PCT/EP200 4 / 0 1 1 5 9 !

Europäisches Patentamt | European Patent Office | Office européen des brevets

REC'D 1 2 NOV 2(
WIPO PC

Bescheinigung | Certificate | Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. | Patent application No. | Demande de brevet n°

03024565.8

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Ofi

Le Président de l'Office européen des breve
p.o.

R C van Dijk

BEST AVAILABLE COPY

Anmeldung Nr:
Application no.: 03024565.8
Demande no:

Anmeldetag:
Date of filing: 28.10.03
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

Bayer HealthCare AG
Kaiser-Wilhelm-Allee
51373 Leverkusen
ALLEMAGNE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Methods and compositions for the response prediction of malignant neoplasia to treatment

In Anspruch genommene Prioriät(en) / Priority(ies) claimed /Priorité(s) revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

C12Q1/68

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL
PT RO SE SI SK TR LI

METHODS AND COMPOSITIONS FOR THE RESPONSE PREDICTION OF MALIGNANT NEOPLASIA TO TREATMENT

TECHNICAL FIELD OF THE INVENTION

The invention relates to methods and compositions for the prediction, diagnosis, prognosis, prevention and treatment of neoplastic disease. Of particular interest is the response prediction of neoplastic lesions to various therapeutic regimens. Neoplastic disease is often caused by chromosomal rearrangements which lead to over- or underexpression of the rearranged genes. The invention discloses genes which are overexpressed in neoplastic tissue and are useful as diagnostic markers and targets for treatment. Methods are disclosed for predicting, diagnosing and prognosing as well as preventing and treating neoplastic disease.

BACKGROUND OF THE INVENTION

Chromosomal aberrations (amplifications, deletions, inversions, insertions, translocations and/or viral integrations) are of importance for the development of cancer and neoplastic lesions, as they account for deregulations of the respective regions. Amplifications of genomic regions have been described, in which genes of importance for growth characteristics, differentiation, invasiveness or resistance to therapeutic intervention are located. One of those regions with chromosomal aberrations is the region carrying the HER-2/neu gene which is amplified in breast cancer patients. In approximately 25% of breast cancer patients the HER-2/neu gene is overexpressed due to gene amplification. HER-2/neu overexpression correlates with a poor prognosis (relapse, overall survival, sensitivity to therapeutics). The importance of HER-2/neu for the prognosis of the disease progression has been described [Gusterson et al., 1992, (1)]. Gene specific antibodies raised against HER-2/neu (Herceptin™) have been generated to treat the respective cancer patients. However, only about 50% of the patients benefit from the antibody treatment with Herceptin™, which is most often combined with chemotherapeutic regimen. The discrepancy of HER-2/neu positive tumors (overexpressing HER-2/neu to similar extent) with regard to responsiveness to therapeutic intervention suggest, that there might be additional factors or genes being involved in growth and apoptotic characteristics of the respective tumor tissues. There seems to be no monocausal relationship between overexpression of the growth factor receptor HER-2/neu and therapy outcome. In line with this the measurement of commonly used tumor markers such as estrogen receptor, progesterone receptor, p53 and Ki-67 do provide only very limited information on clinical outcome of specific therapeutic decisions. Therefore there is a great need for a more detailed diagnostic and prognostic classification of tumors to enable improved therapy decisions and prediction of survival of the patients. The present invention addresses the

need for additional markers by providing genes, which expression is deregulated in tumors and correlates with clinical outcome. One focus is the deregulation of genes present in specific chromosomal regions and their interaction in disease development and drug responsiveness.

HER-2/neu and other markers for neoplastic disease are commonly assayed with diagnostic methods such as immunohistochemistry (IHC) (e.g. HercepTest™ from DAKO Inc.) and Fluorescence-In-Situ-Hybridization (FISH) (e.g. quantitative measurement of the HER-2/neu and Topoisomerase II alpha with a fluorescence-*in-situ*-Hybridization kit from VYSIS). Additionally HER-2/neu can be assayed by detecting HER-2/neu fragments in serum with an ELISA test (BAYER Corp.) or a with a quantitative PCR kit which compares the amount of HER-2/neu gene with the amount of a non-amplified control gene in order to detect HER-2/neu gene amplifications (ROCHE). These methods, however, exhibit multiple disadvantages with regard to sensitivity, specificity, technical and personnel efforts, costs, time consumption, inter-lab reproducibility. These methods are also restricted with regard to measurement of multiple parameters within one patient sample ("multiplexing"). Usually only about 3 to 4 parameters (e.g. genes or gene products) can be detected per tissue slide. Therefore, there is a need to develop a fast and simple test to measure simultaneously multiple parameters in one sample. The present invention addresses the need for a fast and simple high-resolution method, that is able to detect multiple diagnostic and prognostic markers simultaneously.

SUMMARY OF THE INVENTION

The present invention is based on discovery that chromosomal alterations in cancer tissues can lead to changes in the expression of genes that are encoded by the altered chromosomal regions. Exemplary 43 human genes have been identified that are co-amplified in neoplastic lesions from breast cancer tissue resulting in altered expression of several of these genes (Tables 1 to 4). These 43 genes are differentially expressed in breast cancer states, relative to their expression in normal, or non-breast cancer states. The present invention relates to derivatives, fragments, analogues and homologues of these genes and uses or methods of using of the same.

The present invention further relates to novel preventive, predictive, diagnostic, prognostic and therapeutic compositions and uses for malignant neoplasia and breast cancer in particular. Especially membrane bound marker gene products containing extracellular domains can be a particularly useful target for treatment methods as well as diagnostic and clinical monitoring methods.

It is a discovery of the present invention that several of these genes are characterized in that gene products functionally interact in signaling cascades or by directly or indirectly influer each other. This interaction is important for the normal physiology of certain non-neopl tissues (e.g. brain or neurogenic tissue). The deregulation of these genes in neoplastic le: where they are normally exhibit of different level of activity or are not active, however, resul pathophysiology and affects the characteristics of the disease-associated tissue.

The present invention further relates to methods for detecting these deregulations in malig neoplasia on DNA and mRNA level.

The present invention further relates to a method for the detection of chromosomal alterat characterized in that the relative abundance of individual mRNAs, encoded by genes, locate altered chromosomal regions is detected.

The present invention further relates to a method for the detection of the flanking breakpoin named chromosomal alterations by measurement of DNA copy number by quantitative PC DNA-Arrays and DNA sequencing.

A method for the prediction, diagnosis or prognosis of malignant neoplasia by the detectio DNA sequences flanking named genomic breakpoint or are located within such.

The present invention further relates to a method for the detection of chromosomal alterat characterized in that the copy number of one or more genomic nucleic acid sequences loc within an altered chromosomal region(s) is detected by quantitative PCR techniques TaqMan™, Lightcycler™ and iCycler™).

The present invention further relates to a method for the prediction, diagnosis or prognosi malignant neoplasia by the detection of at least 2 markers whereby the markers are genes fragments thereof or genomic nucleic acid sequences that are located on one chromosomal re which is altered in malignant neoplasia and breast cancer in particular.

The present invention also discloses a method for the prediction, diagnosis or prognosi malignant neoplasia by the detection of at least 2 markers whereby the markers are located on or more chromosomal region(s) which is/are altered in malignant neoplasia; and the mar interact as (i) receptor and ligand or (ii) members of the same signal transduction pathwa (iii)members of synergistic signal transduction pathways or (iv) members of antagonistic si transduction pathways or (v) transcription factor and transcription factor binding site.

Also dislcosed is a method for the prediction, diagnosis or prognosis of malignant neoplasia by the detection of at least one marker whereby the marker is a VNTR, SNP, RFLP or STS which is located on one chromosomal region which is altered in malignant neoplasia due to amplification and the marker is detected in (a) a cancerous and (b) a non cancerous tissue or biological sample from the same individual. A preferred embodiment is the detection of at least one VNTR marker of Table 6 or at least on SNP marker of Table 4 or combinations thereof.. Even more preferred can the detection, quantification and sizing of such polymorphic markers be achieved by methods of (a) for the comparative measurement of amount and size by PCR amplification and subsequent capillary electrophoresis, (b) for sequence determination and allelic discrimination by gel electrophoresis (e.g. SSCP, DGGE), real time kinetic PCR, direct DNA sequencing, pyro-sequencing, mass-specific allelic discrimination or resequencing by DNA array technologies, (c) for the dertermination of specific restriction patterns and subsequent electrophoretic separation and (d) for allelic discrimination by allel specific PCR (e.g. ASO). An even more favorable detection of a hetrozygous VNTR, SNP, RFLP or STS is done in a multiplex fashion, utilizing a variety of labeled primers (e.g. fluorescent, radioactive, bioactive) and a suitable capillary electrophoresis (CE) detection system.

In another embodiment the expression of these genes can be detected with DNA-arrays as described in WO9727317 and US6379895.

In a further embodiment the expression of these genes can be detected with bead based direct flourescent readout techniques such as described in WO9714028 and WO9952708.

In one embodiment, the invention pertains to a method of determining the phenotype of a cell or tissue, comprising detecting the differential expression, relative to a normal or untreated cell, of at least one polynucleotide comprising SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19 or 21 to 26 or 53 to 75, wherein the polynucleotide is differentially expressed by at least about 1.5 fold, at least about 2 fold or at least about 3 fold.

In a further aspect the invention pertains to a method of determining the phenotype of a cell or tissue, comprising detecting the differential expression, relative to a normal or untreated cell, of at least one polynucleotide which hybridizes under stringent conditions to one of the polynucleotides of SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19 or 21 to 26 or 53 to 75 and encodes a polypeptide exhibiting the same biological function as given in Table 2 or 3 for the respective polynucleotide, wherein the polynucleotide is differentially expressed by at least at least about 1.5 fold , at least about 2 fold or at least about 3 fold.

In another embodiment of the invention a polynucleotide comprising a polynucleotide select from SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19 or 21 to 26 and 53 to 75 or encoding one of t polypeptides with SEQ ID NO: 28 to 32, 34, 35, 37 to 42, 44, 45 or 47 to 52 or 76 to 98 can used to identify cells or tissue in individuals which exhibit a phenotype predisposed to brea cancer or a diseased phenotype, thereby (a) predicting whether an individual is at risk for tl development, or (b) diagnosing whether an individual is having, or (c) prognosing the progressic or the outcome of the treatment malignant neoplasia and breast cancer in particular.

In yet another embodiment the invention provides a method for identifying genomic regions whic are altered on the chromosomal level and encode genes that are linked by function and ai differentially expressed in malignant neoplasia and breast cancer in particular.

In yet another embodiment the invention provides the genomic regions 17q21, 3p21 and 12q13 fc use in prediction, diagnosis and prognosis as well as prevention and treatment of malignar neoplasia and breast cancer. In particular not only the intragenic regions, but also intergeni regions, pseudogenes or non-transcribed genes of said chromosomal regions can be used fo diagnostic, predictive, prognostic and preventive and therapeutic compositions and methods Therefore sequences of coding or non-coding regions as depicted in this invention are offered by way of illustration and not by way of limitation. As one aspect of this, genomic sequences ir between the genomic sequences depicted can be used for similiar purposes.

In yet another embodiment the invention provides methods of screening for agents which regulate the activity of a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75. A test compound is contacted with a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75. Binding of the test compound to the polypeptide is detected. A test compound which binds to the polypeptide is thereby identified as a potential therapeutic agent for the treatment of malignant neoplasia and more particularly breast cancer.

In even another embodiment the invention provides another method of screening for agents which regulate the activity of a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75. A test compound is contacted with a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75. A biological activity

mediated by the polypeptide is detected. A test compound which decreases the biological activity is thereby identified as a potential therapeutic agent for decreasing the activity of the polypeptide encoded by a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 in malignant neoplasia and breast cancer in particular. A test compound which increases the biological activity is thereby identified as a potential therapeutic agent for increasing the activity of the polypeptide encoded by a polypeptide selected from one of the polypeptides with SEQ ID NO: 27 to 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 in malignant neoplasia and breast cancer in particular.

In another embodiment the invention provides a method of screening for agents which regulate the activity of a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75. A test compound is contacted with a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75. Binding of the test compound to the polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 is detected. A test compound which binds to the polynucleotide is thereby identified as a potential therapeutic agent for regulating the activity of a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 in malignant neoplasia and breast cancer in particular.

The invention thus provides polypeptides selected from one of the polypeptides with SEQ ID NO: 27 to 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 which can be used to identify compounds which may act, for example, as regulators or modulators such as agonists and antagonists, partial agonists, inverse agonists, activators, co-activators and inhibitors of the polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75. Accordingly, the invention provides reagents and methods for regulating a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 in malignant neoplasia and more particularly breast cancer. The regulation can be an up- or down regulation. Reagents that modulate the expression, stability or amount of a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 or the activity of the polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 can be a protein, a peptide, a peptidomimetic, a nucleic acid, a nucleic acid analogue (e.g. peptide nucleic acid,

locked nucleic acid) or a small molecule. Methods that modulate the expression, stability amount of a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and : to 75 or the activity of the polypeptide comprising a polypeptide selected from SEQ ID NO: 27 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SE ID NO: 1 to 26 and 53 to 75 can be gene replacement therapies, antisense, ribozyme and triple nucleic acid approaches.

In one embodiment of the invention provides antibodies which specifically bind to a full-length (partial polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 (encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 an 53 to 75 or a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 an 53 to 75 for use in prediction, prevention, diagnosis, prognosis and treatment of malignar neoplasia and breast cancer in particular.

Yet another embodiment of the invention is the use of a reagent which specifically binds to : polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 or : polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or encode(by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 7: in the preparation of a medicament for the treatment of malignant neoplasia and breast cancer ir particular.

Still another embodiment is the use of a reagent that modulates the activity or stability of a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 or the expression, amount or stability of a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 in the preparation of a medicament for the treatment of malignant neoplasia and breast cancer in particular.

Still another embodiment of the invention is a pharmaceutical composition which includes a reagent which specifically binds to a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 or a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75, and a pharmaceutical acceptable carrier.

Yet another embodiment of the invention is a pharmaceutical composition including a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 or

encoding a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98.

In one embodiment, a reagent which alters the level of expression in a cell of a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 or encoding a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98, or a sequence complementary thereto, is identified by providing a cell, treating the cell with a test reagent, determining the level of expression in the cell of a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 or encoding a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or a sequence complementary thereto, and comparing the level of expression of the polynucleotide in the treated cell with the level of expression of the polynucleotide in an untreated cell, wherein a change in the level of expression of the polynucleotide in the treated cell relative to the level of expression of the polynucleotide in the untreated cell is indicative of an agent which alters the level of expression of the polynucleotide in a cell.

The invention further provides a pharmaceutical composition comprising a reagent identified by this method.

Another embodiment of the invention is a pharmaceutical composition which includes a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98 or which is encoded by a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75.

A further embodiment of the invention is a pharmaceutical composition comprising a polynucleotide including a sequence which hybridizes under stringent conditions to a polynucleotide comprising a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 and encoding a polypeptide exhibiting the same biological function as given for the respective polynucleotide in Table 2 or 3, or encoding a polypeptide comprising a polypeptide selected from SEQ ID NO: 27 to 52 and 76 to 98. Pharmaceutical compositions, useful in the present invention may further include fusion proteins comprising a polypeptide comprising a polynucleotide selected from SEQ ID NO: 27 to 52 and 76 to 98, or a fragment thereof, antibodies, or antibody fragments

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows a sketch of the chromosome 17 with G-banding pattern and cytogenetic positions. In the blow out at the lower part of the figure a detailed view of the chromosomal area of the long arm of chromosome 17 (17q12-21.1) is provided. Each vertical rectangle depicted

in medium gray, represents a gene as labeled below or above the individual position. The order of genes depicted in this graph has been deduced from experiments questioning the amplification an over expression and from public available data (e.g. UCSC, NCBI or Ensemble).

Fig. 2 shows the same region as depicted before in Fig. 1 and a cluster representation of the individual expression values measured by DNA-chip hybridization. The gene representing squares are indicated by a dotted line. In the upper part of the cluster representation 4 tumor cell lines, of which two harbor a known HER-2/neu over expression (SKBR3 and AU565), are depicted with their individual expression profiles. Not only the HER-2/neu gene shows a clear over expression but as provided by this invention several other genes with in the surrounding. In the middle part of the cluster representation expression data obtained from immune histochemically characterized tumor samples are presented. Two of the depicted probes show a significant over expression of genes marked by the white rectangles. For additional information and comparison expression profiles of several non diseased human tissues (rnas obtained from Clontech Inc.) Are provided. Closest relation to the expression profile of HER-2/neu positive tumors displays human brain and neural tissue.

Fig. 3 provides data from DNA amplification measurements by qpcr (e.g. Taqman). Data indicates that in several analyzed breast cancer cell lines harbor amplification of genes which were located in the previously described region (ARCHEON). Data were displayed for each gene on the x-axis and 40-Ct at the y-axis. Data were normalized to the expression level of GAPDH as seen in the first group of columns.

Fig. 4 represents a graphical overview on the amplified regions and provides information on the length of the individual amplification and over expression in the analyzed tumor cell lines. The length of the amplification and the composition of genes has a significant impact on the nature of the cancer cell and on the responsiveness on certain drugs, as described elsewhere.

DETAILED DESCRIPTION OF THE INVENTION

DEFINITIONS

"Differential expression", as used herein, refers to both quantitative as well as qualitative differences in the genes' expression patterns depending on differential development and/or tumor growth. Differentially expressed genes may represent "marker genes," and/or "target genes". The

expression pattern of a differentially expressed gene disclosed herein may be utilized as part of a prognostic or diagnostic breast cancer evaluation. Alternatively, a differentially expressed gene disclosed herein may be used in methods for identifying reagents and compounds and uses of these reagents and compounds for the treatment of breast cancer as well as methods of treatment.

"Biological activity" or "bioactivity" or "activity" or "biological function", which are used interchangeably, herein mean an effector or antigenic function that is directly or indirectly performed by a polypeptide (whether in its native or denatured conformation), or by any fragment thereof *in vivo* or *in vitro*. Biological activities include but are not limited to binding to polypeptides, binding to other proteins or molecules, enzymatic activity, signal transduction, activity as a DNA binding protein, as a transcription regulator, ability to bind damaged DNA, etc. A bioactivity can be modulated by directly affecting the subject polypeptide. Alternatively, a bioactivity can be altered by modulating the level of the polypeptide, such as by modulatin expression of the corresponding gene.

The term "marker" or "biomarker" refers a biological molecule, e.g., a nucleic acid, peptide, hormone, etc., whose presence or concentration can be detected and correlated with a known condition, such as a disease state.

"Marker gene," as used herein, refers to a differentially expressed gene which expression pattern may be utilized as part of predictive, prognostic or diagnostic malignant neoplasia or breast cancer evaluation, or which, alternatively, may be used in methods for identifying compounds useful for the treatment or prevention of malignant neoplasia and breast cancer in particular. A marker gene may also have the characteristics of a target gene.

"Target gene", as used herein, refers to a differentially expressed gene involved in breast cancer a manner by which modulation of the level of target gene expression or of target gene product activity may act to ameliorate symptoms of malignant neoplasia and breast cancer in particular. A target gene may also have the characteristics of a marker gene.

The term "biological sample", as used herein, refers to a sample obtained from an organism or from components (e.g., cells) of an organism. The sample may be of any biological tissue or fluid. Frequently the sample will be a "clinical sample" which is a sample derived from a patient. Such samples include, but are not limited to, sputum, blood, blood cells (e.g., white cells), tissue or fine needle biopsy samples, cell-containing bodyfluids, free floating nucleic acids, urine, peritoneal fluid, and pleural fluid, or cells therefrom. Biological samples may also include sections of tissues such as frozen sections taken for histological purposes.

By "array" or "matrix" is meant an arrangement of addressable locations or "addresses" on device. The locations can be arranged in two dimensional arrays, three dimensional arrays, other matrix formats. The number of locations can range from several to at least hundreds thousands. Most importantly, each location represents a totally independent reaction site. Arra include but are not limited to nucleic acid arrays, protein arrays and antibody arrays. A "nucle acid array" refers to an array containing nucleic acid probes, such as oligonucleotide polynucleotides or larger portions of genes. The nucleic acid on the array is preferably sing stranded. Arrays wherein the probes are oligonucleotides are referred to as "oligonucleotic arrays" or "oligonucleotide chips." A "microarray," herein also refers to a "biochip" or "biologic chip", an array of regions having a density of discrete regions of at least about 100/cm^2, an preferably at least about 1000/cm^2. The regions in a microarray have typical dimensions, e.g diameters, in the range of between about 10-250 μm, and are separated from other regions in th array by about the same distance. A "protein array" refers to an array containing polypeptid probes or protein probes which can be in native form or denatured. An "antibody array" refers t an array containing antibodies which include but are not limited to monoclonal antibodies (e.g from a mouse), chimeric antibodies, humanized antibodies or phage antibodies and single chai antibodies as well as fragments from antibodies.

The term "agonist", as used herein, is meant to refer to an agent that mimics or upregulates (e.g., potentiates or supplements) the bioactivity of a protein. An agonist can be a wild-type protein or derivative thereof having at least one bioactivity of the wild-type protein. An agonist can also be a compound that upregulates expression of a gene or which increases at least one bioactivity of a protein. An agonist can also be a compound which increases the interaction of a polypeptide with another molecule, e.g., a target peptide or nucleic acid.

The term "antagonist" as used herein is meant to refer to an agent that downregulates (e.g., suppresses or inhibits) at least one bioactivity of a protein. An antagonist can be a compound which inhibits or decreases the interaction between a protein and another molecule, e.g., a target peptide, a ligand or an enzyme substrate. An antagonist can also be a compound that downregulates expression of a gene or which reduces the amount of expressed protein present.

"Small molecule" as used herein, is meant to refer to a composition, which has a molecular weight of less than about 5 kD and most preferably less than about 4 kD. Small molecules can be nucleic acids, peptides, polypeptides, peptidomimetics, carbohydrates, lipids or other organic (carbon-containing) or inorganic molecules. Many pharmaceutical companies have extensive libraries of chemical and/or biological mixtures, often fungal, bacterial, or algal extracts, which can be screened with any of the assays of the invention to identify compounds that modulate a bioactivity.

The terms “modulated" or "modulation" or "regulated" or "regulation" and “differentially regulated” as used herein refer to both upregulation (i.e., activation or stimulation (e.g., by agonizing or potentiating) and down regulation [i.e., inhibition or suppression (e.g., by antagonizing, decreasing or inhibiting)].

“Transcriptional regulatory unit” refers to DNA sequences, such as initiation signals, enhancers, and promoters, which induce or control transcription of protein coding sequences with which they are operably linked. In preferred embodiments, transcription of one of the genes is under the control of a promoter sequence (or other transcriptional regulatory sequence) which controls the expression of the recombinant gene in a cell-type in which expression is intended. It will also be understood that the recombinant gene can be under the control of transcriptional regulatory sequences which are the same or which are different from those sequences which control transcription of the naturally occurring forms of the polypeptide.

The term “derivative” refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

The term “nucleotide analog” refers to oligomers or polymers being at least in one feature different from naturally occurring nucleotides, oligonucleotides or polynucleotides, but exhibiting functional features of the respective naturally occurring nucleotides (e.g. base paring, hybridization, coding information) and that can be used for said compositions. The nucleotide analogs can consist of non-naturally occurring bases or polymer backbones, examples of which are LNAs, PNAs and Morpholinos. The nucleotide analog has at least one molecule different from its naturally occurring counterpart or equivalent.

"BREAST CANCER GENES" or "BREAST CANCER GENE" as used herein refers to the polynucleotides of SEQ ID NO: 1 to 26 and 53 to 75, as well as derivatives, fragments, analogs and homologues thereof, the polypeptides encoded thereby, the polypeptides of SEQ ID NO: 27 to 52 and 76 to 98 as well as derivatives, fragments, analogs and homologues thereof and the corresponding genomic transcription units which can be derived or identified with standard techniques well known in the art using the information disclosed in Tables 1 to 5 and Figures 1 to 4. The GenBank, Locuslink ID and the UniGene accession numbers of the polynucleotide

sequences of the SEQ ID NO: 1 to 26 and 53 to 75 and the polypeptides of the SEQ ID NO: 2 52 and 76 to 98 are shown in Table 1, the gene description, gene function and subcell localization is given in Tables 2 and 3.

The term "chromosomal region" as used herein refers to a consecutive DNA stretch c chromosome which can be defined by cytogenetic or other genetic markers such as e.g. restric length polymorphisms (RFLPs), single nucleotide polymorphisms (SNPs), expressed sequ tags (ESTs), sequence tagged sites (STSs), microsatellites, variable number of tandem rep (VNTRs) and genes. Typically a chromosomal region consists of up to 2 Megabases (MB), up MB, up to 6 MB, up to 8 MB, up to 10 MB, up to 20 MB or even more MB.

The term "altered chromosomal region" or" abberant chromosomal region" refers to a struct change of the chromosomal composition and DNA sequence, which can occur by the follov events: amplifications, deletions, inversions, insertions, translocations and/or viral integration. trisomy, where a given cell harbors more than two copies of a chromosome, is within the mea of the term "amplification" of a chromosome or chromosomal region.

The present invention provides polynucleotide sequences and proteins encoded thereby, as we probes derived from the polynucleotide sequences, antibodies directed to the encoded proteins, predictive, preventive, diagnostic, prognostic and therapeutic uses for individuals which are at for or which have malignant neoplasia and breast cancer in particular. The sequences disclo herein have been found to be differentially expressed in samples from breast cancer.

The present invention is based on the identification of 43 genes that are differentially regul (up- or downregulated) in tumor biopsies of patients with clinical evidence of breast cancer. identification of 43 human genes which were not known to be differentially regulated in br cancer states and their significance for the disease is described in the working examples her The characterization of the co-expression of these genes provides newly identified roles in br cancer. The gene names, the database accession numbers (GenBank and UniGene) as well as putative or known functions of the encoded proteins and their subcellular localization are give Tables 1 to 4. The primer sequences used for the gene amplification are shown in Table 5.

In either situation, detecting expression of these genes in excess or in with lower leve compared to normal expression provides the basis for the diagnosis of malignant neoplasia breast cancer. Furthermore, in testing the efficacy of compounds during clinical trials, a decr in the level of the expression of these genes corresponds to a return from a disease condition normal state, and thereby indicates a positive effect of the compound.

Another aspect of the present invention is based on the observation that neighboring genes within defined genomic regions functionally interact and influence each others function directly or indirectly. A genomic region encoding functionally interacting genes that are co-amplified and co-expressed in neoplastic lesions has been defined as an "ARCHEON". (ARCHEON = Altered Region of Changed Chromosomal Expression Observed in Neoplasms). Chromosomal alterations often affect more than one gene. This is true for amplifications, duplications, insertions, integrations, inversions, translocations, and deletions. These changes can have influence on the expression level of single or multiple genes. Most commonly in the field of cancer diagnostics and treatment the changes of expression levels have been investigated for single, putative relevant target genes such as MLVI2 (5p14), NRASL3 (6p12), EGFR (7p12), c-myc (8q23), Cyclin D1 (11q13), IGF1R (15q25), HER-2/neu (17q21), PCNA (20q12). However, the altered expression level and interaction of multiple (i.e. more than two) genes within one genomic region with each other has not been addressed. Genes of an ARCHEON form gene clusters with tissue specific expression patterns. The mode of interaction of individual genes within such a gene cluster suspected to represent an ARCHEON can be either protein-protein or protein-nucleic acid interaction, which may be illustrated but not limited by the following examples: ARCHEON gene interaction may be in the same signal transduction pathway, may be receptor to ligand binding, receptor kinase and SH2 or SH3 binding, transcription factor to promoter binding, nuclear hormone receptor to transcription factor binding, phosphogroup donation (e.g. kinases) and acceptance (e.g. phosphoprotein), mRNA stabilizing protein binding and transcriptional processes. The individual activity and specificity of a pair genes and or the proteins encoded thereby or of a group of such in a higher order, may be readily deduced from literature, published or deposited within public databases by the skilled person. However in the context of an ARCHEON the interaction of members being part of an ARCHEON will potentiate, exaggerate or reduce their singular functions. This interaction is of importance in defined normal tissues in which they are normally co-expressed. Therefore, these clusters have been commonly conserved during evolution. The aberrant expression of members of these ARCHEON in neoplastic lesions, however, (especially within tissues in which they are normally not expressed) has influence on tumor characteristics such as growth, invasiveness and drug responsiveness. Due to the interaction of these neighboring genes it is of importance to determine the members of the ARCHEON which are involved in the deregulation events. In this regard amplification and deletion events in neoplastic lesions are of special interest.

The invention relates to a method for the detection of chromosomal alterations by (a) determining the relative mRNA abundance of individual mRNA species or (b) determining the copy number of one or more chromosomal region(s) by quantitative PCR. In one embodiment information on the

genomic organization and spatial regulation of chromosomal regions is assessed by bioinforn analysis of the sequence information of the human genome (UCSC, NCBI) and then comb with RNA expression data from GeneChip™ DNA-Arrays (Affymetrix) and/or quantitative : (TaqMan) from RNA-samples or genomic DNA.

In a further embodiment the functional relationship of genes located on a chromosomal re which is altered (amplified or deleted) is established. The altered chromosomal region is def as an ARCHEON if genes located on that region functionally interact.

The 17q21 locus was investigated as one model system, harboring the HER-2/neu gene. establishing a high-resolution assay to detect amplification events in neighboring genes, 43 g that are commonly co-amplified in breast cancer cell lines and patient samples were identified gene array technologies and immunological methods their co-overexpression in tumor sam was demonstrated. Surprisingly, by clustering tissue samples with HER-2/neu positive Tu samples, it was found that the expression pattern of this larger genomic region (consisting o genes) is very similar to control brain tissue. HER-2/neu negative breast tumor tissue did not s a similar expression pattern. Indeed, some of the genes within these cluster are important neural development (HER-2/neu, THRA) in mouse model systems or are described tc expressed in neural cells (NeuroD2). Moreover, by searching similar gene combinations in human and rodent genome additional homologous chromosomal regions on chromosome 3p21 12q13 harboring several isoforms of the respective genes (see below) were found. There w strong evidence for multiple interactions between the 43 candidate genes, as being part of ident pathways (HER-2, neu, GRB7, CrkRS, CDC6), influencing the expression of each other (H 2/neu, THRA, RARA), interacting with each other (PPARGBP, THRA, RARA, NR1D1 or H 2/neu, GRB7) or expressed in defined tissues (CACNB1, PPARGBP, etc.). Interestingly, genomic regions of the ARCHEONs that were identified are amplified in acquired Tamox resistance of HER-2/neu negative cells (MCF7), which are normally sensitive to Tamox treatment [Achuthan et al., 2001,(2)].

Moreover, altered responsiveness to treatment due to the alterations of the genes within tl ARCHEONs was observed. Surprisingly, genes within the ARCHEONs are of importance eve the absence of HER-2/neu homologues. Some of the genes within the ARCHEONs, do not (serve as marker genes for prognostic purposes, but have already been known as targets therapeutic intervention. For example TOP2 alpha is a target of anthracyclins. THRA and RA can be targeted by hormones and hormone analogs (e.g. T3, rT3, RA). Due to their high affi binding sites and available screening assays (reporter assays based on their transcriptio potential) the hormone receptors which are shown to be linked to neoplastic pathophysiology

the first time herein are ideal targets for drug screening and treatment of malignant neoplasia and breast cancer in particular. In this regard it is essential to know which members of the ARCHEON are altered in the neoplastic lesions. Particularly it is important to know the nature, number and extent to which the ARCHEON genes are amplified or deleted. The ARCHEONs are flanked by similar, endogenous retroviruses (e.g. HERV-K= "human endogenous retrovirus"), some of which are activated in breast cancer. These viruses may have also been involved in the evolutionary duplication of the ARCHEONs.

The analysis of the 17q21 region proved data obtained by IHC and identified several additional genes being co-amplified with the HER-2/neu gene. Comparative Analysis of RNA-based quantitative RT-PCR (TaqMan) with DNA-based qPCR from tumor cell lines identified the same amplified region. Genes at the 17q11.2 –21. region are offered by way of illustration not by way of limitation. A graphical display of the described chromosomal region is provided in Figure 1.

Biological relevance of the genes which are part of the 17q21 ARCHEON

MLN50

By differential screening of cDNAs from breast cancer-derived metastatic axillary lymph nodes, TRAF4 and 3 other novel genes (MLN51, MLN62, MLN64) were identified that are overexpressed in breast cancer [Tomasetto et al., 1995, (3)]. One gene, which they designated MLN50, was mapped to 17q11-q21.3 by radioactive in situ hybridization. In breast cancer cell lines, overexpression of the 4 kb MLN50 mRNA was correlated with amplification of the gene and with amplification and overexpression of ERBB2, which maps to the same region. The authors suggested that the 2 genes belong to the same amplicon. Amplification of chromosomal region 17q11-q21 is one of the most common events occurring in human breast cancers. They report that the predicted 261-amino acid MLN50 protein contains an N-terminal LIM domain and a C-terminal SH3 domain. They renamed the protein LASP1, for 'LIM and SH3 protein.' Northern blot analysis revealed that LASP1 mRNA was expressed at a basal level in all normal tissues examined and overexpressed in 8% of primary breast cancers. In most of these cancers, LASP1 and ERBB2 were simultaneously overexpressed.

MLLT6

The MLLT6 (AF17) gene encodes a protein of 1,093 amino acids, containing a leucine-zipper dimerization motif located 3-prime of the fusion point and a cysteine-rich domain at the end terminus. AF17 was found to contain stretches of amino acids previously associated with domains involved in transcriptional repression or activation.

Chromosome translocations involving band 11q23 are associated with approximately 10 patients with acute lymphoblastic leukemia (ALL) and more than 5% of patients with myeloid leukemia (AML). The gene at 11q23 involved in the translocations is vari designated ALL1, HRX, MLL, and TRX1. The partner gene in one of the rarer transloca t(11;17)(q23;q21), designated MLLT6 on 17q12.

ZNF144 (Mel18)

Mel18 cDNA encodes a novel cys-rich zinc finger motif. The gene is expressed strongly in tumor cell lines, but its normal tissue expression was limited to cells of neural origin and especially abundant in fetal neural cells. It belongs to a RING-finger motif family which inc BMI1. The MEL18/BMI1 gene family represents a mammalian homolog of the Droso 'polycomb' gene group, thereby belonging to a memory mechanism involved in maintainin the expression pattern of key regulatory factors such as Hox genes. Bmi1, Mel18 and M33 g as representative examples of mouse Pc-G genes. Common phenotypes observed in knockout mutant for each of these genes indicate an important role for Pc-G genes not only in regulati Hox gene expression and axial skeleton development but also in control of proliferation survival of haematopoietic cell lineages. This is in line with the observed prolifer deregulation observed in lymphoblastic leukemia. The MEL18 gene is conserved ar vertebrates. Its mRNA is expressed at high levels in placenta, lung, and kidney, and at lower l in liver, pancreas, and skeletal muscle. Interestingly, cervical and lumbo-sacral-HOX expression is altered in several primary breast cancers with respect to normal breast tissue wit HoxB gene cluster being present on 17q distal to the 17q21 locus. Moreover, dela differentiation with persistent nests of proliferating cells was found in endothelial cells cocult with HOXB7-transduced SkBr3 cells, which exhibit a 17q21 amplification. Tumorigenicit these cells has been evaluated in vivo. Xenograft in athymic nude mice showed SkBr3/HOXB7 cells developed tumors with an increased number of blood vessels, e irradiated or not, whereas parental SkBr3 cells did not show any tumor take unless mice sublethally irradiated. As part of this invention, we have found MEL18 to be overexpre specifically in tumors bearing Her-2/neu gene amplification, which can be critical for expression.

PHOSPHATIDYLINOSITOL-4-PHOSPHATE 5-KINASE, TYPE II, BETA; PIP5K2B

Phosphoinositide kinases play central roles in signal transduction. Phosphatidylinosit phosphate 5-kinases (PIP5Ks) phosphorylate phosphatidylinositol 4-phosphate, giving ris phosphatidylinositol 4,5-bisphosphate. The PIP5K enzymes exist as multiple isoforms that

various immunoreactivities, kinetic properties, and molecular masses. They are unique in that they possess almost no homology to the kinase motifs present in other phosphatidylinositol, protein, and lipid kinases. By screening a human fetal brain cDNA library with the PIP5K2B EST the full length gene could be isolated. The deduced 416-amino acid protein is 78% identical to PIP5K2A. Using SDS-PAGE, the authors estimated that bacterially expressed PIP5K2B has a molecular mass of 47 kD. Northern blot analysis detected a 6.3-kb PIP5K2B transcript which was abundantly expressed in several human tissues. PIP5K2B interacts specifically with the juxtamembrane region of the p55 TNF receptor (TNFR1) and PIP5K2B activity is increased in mammalian cells by treatment with TNF-alpha. A modeled complex with membrane-bound substrate and ATP shows how a phosphoinositide kinase can phosphorylate its substrate in situ at the membrane interface.The substrate-binding site is open on 1 side, consistent with dual specificity for phosphatidylinositol 3- and 5-phosphates. Although the amino acid sequence of PIP5K2A does not show homology to known kinases, recombinant PIP5K2A exhibited kinase activity. PIP5K2A contains a putative Src homology 3 (SH3) domain-binding sequence. Overexpression of mouse PIP5K1B in COS7 cells induced an increase in short actin fibers and a decrease in actin stress fibers.

TEM7

Using serial analysis of gene expression (SAGE) a partial cDNAs corresponding to several tumor endothelial markers (TEMs) that displayed elevated expression during tumor angiogenesis could be identified. Among the genes identified was TEM7. Using database searches and 5-prime RACE the entire TEM7 coding region, which encodes a 500-amino acid type I transmembrane protein,has been described.. The extracellular region of TEM7 contains a plexin-like domain and has weak homology to the ECM protein nidogen. The function of these domains, which are usually found in secreted and extracellular matrix molecules, is unknown. Nidogen itself belongs to the entactin protein family and helps to determine pathways of migrating axons by switching from circumferential to longitudinal migration. Entactin is involved in cell migration, as it promotes trophoblast outgrowth through a mechanism mediated by the RGD recognition site, and plays an important role during invasion of the endometrial basement membrane at implantation. As entactin promotes thymocyte adhesion but affects thymocyte migration only marginally, it is suggested that entactin may plays a role in thymocyte localization during T cell development.

In situ hybridization analysis of human colorectal cancer demonstrated that TEM7 was expressed clearly in the endothelial cells of the tumor stroma but not in the endothelial cells of normal colonic tissue. Using in situ hybridization to assay expression in various normal adult mouse

tissues, they observed that TEM7 was largely undetectable in mouse tissues or tumors, but was abundantly expressed in mouse brain.

ZNFN1A3

By screening a B-cell cDNA library with a mouse Aiolos N-terminal cDNA probe, a cDNA encoding human Aiolos, or ZNFN1A3, was obtained. The deduced 509-amino acid protein, which is 86% identical to its mouse counterpart, has 4 DNA-binding zinc fingers in its N terminus and 2 zinc fingers that mediate protein dimerization in its C terminus. These domains are 100% and 96% homologous to the corresponding domains in the mouse protein, respectively. Northern blot analysis revealed strong expression of a major 11.0- and a minor 4.4-kb ZNFN1A3 transcript in peripheral blood leukocytes, spleen, and thymus, with lower expression in liver, small intestine, and lung.

Ikaros (ZNFN1A1), a hemopoietic zinc finger DNA-binding protein, is a central regulator of lymphoid differentiation and is implicated in leukemogenesis. The execution of normal function of Ikaros requires sequence-specific DNA binding, transactivation, and dimerization domains. Mice with a mutation in a related zinc finger protein, Aiolos, are prone to B-cell lymphoma. In chemically induced murine lymphomas allelic losses on markers surrounding the Znfn1a1 gene were detected in 27% of the tumors analyzed. Moreover specific Ikaros expression was in primary mouse hormone-producing anterior pituitary cells and substantial for Fibroblast growth factor receptor 4 (FGFR4) expression, which itself is implicated in a multitude of endocrine cell hormonal and proliferative properties with FGFR4 being differentially expressed in normal and neoplastic pituitary. Moreover Ikaros binds to chromatin remodelling complexes containing SWI/SNF proteins, which antagonize Polycomb function. Intetrestingly at the telomeric end of the disclosed ARCHEON the SWI/SNF complex member SMARCE1 (= SWI/SNF-related, matrix-associated, actin-dependent regulators of chromatin) is located and part of the described amplification. Due to the related binding specificities of Ikaros and Palindrom Binding Protein (PBP) it is suggestive, that ZNFN1A3 is able to regulate the Her-2/neu enhancer.

PPP1R1B

Midbrain dopaminergic neurons play a critical role in multiple brain functions, and abnormal signaling through dopaminergic pathways has been implicated in several major neurologic and psychiatric disorders. One well-studied target for the actions of dopamine is DARPP32. In the densely dopamine- and glutamate-innervated rat caudate-putamen, DARPP32 is expressed in medium-sized spiny neurons that also express dopamine D1 receptors. The function of DARPP32

seems to be regulated by receptor stimulation. Both dopaminergic and glutamatergic (NMDA) receptor stimulation regulate the extent of DARPP32 phosphorylation, but in opposite directions.

The human DARPP32 was isolated from a striatal cDNA library. The 204-amino acid DARPP32 protein shares 88% and 85% sequence identity, respectively, with bovine and rat DARPP32 proteins. The DARPP32 sequence is particularly conserved through the N terminus, which represents the active portion of the protein. Northern blot analysis demonstrated that the 2.1-kb DARPP32 mRNA is more highly expressed in human caudate than in cortex. In situ hybridization to postmortem human brain showed a low level of DARPP32 expression in all neocortical layers, with the strongest hybridization in the superficial layers. CDK5 phosphorylated DARPP32 in vitro and in intact brain cells. Phospho-thr75 DARPP32 inhibits PKA in vitro by a competitive mechanism. Decreasing phospho-thr75 DARPP32 in striatal cells either by a CDK5-specific inhibitor or by using genetically altered mice resulted in increased dopamine-induced phosphorylation of PKA substrates and augmented peak voltage-gated calcium currents. Thus, DARPP32 is a bifunctional signal transduction molecule which, by distinct mechanisms, controls a serine/threonine kinase and a serine/threonine phosphatase.

DARPP32 and t-DARPP are overexpressed in gastric cancers. It's suggested that overexpression of these 2 proteins in gastric cancers may provide an important survival advantage to neoplastic cells. It could be demonstrated that Darpp32 is an obligate intermediate in progesterone-facilitated sexual receptivity in female rats and mice. The facilitative effect of progesterone on sexual receptivity in female rats was blocked by antisense oligonucleotides to Darpp32. Homozygous mice carrying a null mutation for the Darpp32 gene exhibited minimal levels of progesterone-facilitated sexual receptivity when compared to their wildtype littermates, and progesterone significantly increased hypothalamic cAMP levels and cAMP-dependent protein kinase activity.

CACNB1

In 1991 a cDNA clone encoding a protein with high homology to the beta subunit of the rabbit skeletal muscle dihydropyridine-sensitive calcium channel from a rat brain cDNA library [Pragnell et al., 1991, (4)]. This rat brain beta-subunit cDNA hybridized to a 3.4-kb message that was expressed in high levels in the cerebral hemispheres and hippocampus and much lower levels in cerebellum. The open reading frame encodes 597 amino acids with a predicted mass of 65,679 Da which is 82% homologous with the skeletal muscle beta subunit. The corresponding human beta-subunit gene was localized to chromosome 17 by analysis of somatic cell hybrids. The authors suggested that the encoded brain beta subunit, which has a primary structure highly similar to its

isoform in skeletal muscle, may have a comparable role as an integral regulatory component oi neuronal calcium channel.

RPL19

The ribosome is the only organelle conserved between prokaryotes and eukaryotes. In eukaryote this organelle consists of a 60S large subunit and a 40S small subunit. The mammalian ribosom contains 4 species of RNA and approximately 80 different ribosomal proteins, most of whic appear to be present in equimolar amounts. In mammalian cells, ribosomal proteins can accou for up to 15% of the total cellular protein, and the expression of the different ribosomal protei genes, which can account for up to 7 to 9% of the total cellular mRNAs, is coordinately regulate to meet the cell's varying requirements for protein synthesis. The mammalian ribosomal protei genes are members of multigene families, most of which are composed of multiple processe pseudogenes and a single functional intron-containing gene. The presence of multiple pseudogene hampered the isolation and study of the functional ribosomal protein genes. By study of somati cell hybrids, it has been elucidated that DNA sequences complementary to 6 mammalia ribosomal protein cDNAs could be assigned to chromosomes 5, 8, and 17. Ten fragments mappe to 3 chromosomes [Nakamichi et al., 1986, (5)]. These are probably a mixture of functiona (expressed) genes and pseudogenes. One that maps to 5q23-q33 rescues Chinese hamster emetine-resistance mutations in interspecies hybrids and is therefore the transcriptionally active RPS14 gene. In 1989 a PCR-based strategy for the detection of intron-containing genes in the presence of multiple pseudogenes was described. This technique was used to identify the intron-containing PCR products of 7 human ribosomal protein genes and to map their chromosomal locations by hybridization to human/rodent somatic cell hybrids [Feo et al., 1992, (6)]. All 7 ribosomal protein genes were found to be on different chromosomes: RPL19 on 17p12-q11;RPL30 on 8; RPL35A on 18; RPL36A on 14; RPS6 on 9pter-p13; RPS11 on 19cen-qter; and RPS17 on 11pter-p13. These are also different sites from the chromosomal location of previously mapped ribosomal protein genes S14 on chromosome 5, S4 on Xq and Yp, and RP117A on 9q3-q34. By fluorescence in situ hybridization the position of the RPL19 gene was mapped to 17q11 [Davies et al., 1989, (7)].

PPARBP, PBP, CRSP1, CRSP200, TRIP2, TRAP220, RB18A, DRIP230

The thyroid hormone receptors (TRs) are hormone-dependent transcription factors that regulate expression of a variety of specific target genes. They must specifically interact with a number of proteins as they progress from their initial translation and nuclear translocation to heterodimerization with retinoid X receptors (RXRs), functional interactions with other transcription factors and the basic transcriptional apparatus, and eventually, degradation. To help

elucidate the mechanisms that underlie the transcriptional effects and other potential functions of TRs, the yeast interaction trap, a version of the yeast 2-hybrid system, was used to identify proteins that specifically interact with the ligand-binding domain of rat TR-beta-1 (THRB) [Lee et al., 1995, (8)]. The authors isolated HeLa cell cDNAs encoding several different TR-interacting proteins (TRIPs), including TRIP2. TRIP2 interacted with rat Thrb only in the presence of thyroid hormone. It showed a ligand-independent interaction with RXR-alpha, but did not interact with the glucocorticoid receptor (NR3C1) under any condition. By immunoscreening a human B-lymphoma cell cDNA expression library with the anti-p53 monoclonal antibody PAb1801, PPARBP was identified, which was called RB18A for 'recognized by PAb1801 monoclonal antibody' [Drane et al., 1997, (9)]. The predicted 1,566-amino acid RB18A protein contains several potential nuclear localization signals, 13 potential N-glycosylation sites, and a high number of potential phosphorylation sites. Despite sharing common antigenic determinants with p53, RB18A does not show significant nucleotide or amino acid sequence similarity with p53. Whereas the calculated molecular mass of RB18A is 166 kD, the apparent mass of recombinant RB18A was 205 kD by SDS-PAGE analysis. The authors demonstrated that RB18A shares functional properties with p53, including DNA binding, p53 binding, and self-oligomerization. Furthermore, RB18A was able to activate the sequence-specific binding of p53 to DNA, which was induced through an unstable interaction between both proteins. Northern blot analysis of human tissues detected an 8.5-kb RB18A transcript in all tissues examined except kidney, with highest expression in heart. Moreover mouse Pparbp, which was called Pbp for 'Ppar-binding protein,' as a protein that interacts with the Ppar-gamma (PPARG) ligand-binding domain in a yeast 2-hybrid system was identified [Zhu et al., 1997, (10)]. The authors found that Pbp also binds to PPAR-alpha (PPARA), RAR-alpha (RARA), RXR, and TR-beta-1 in vitro. The binding of Pbp to these receptors increased in the presence of specific ligands. Deletion of the last 12 amino acids from the C terminus of PPAR-gamma resulted in the abolition of interaction between Pbp and PPAR-gamma. Pbp modestly increased the transcriptional activity of PPAR-gamma, and a truncated form of Pbp acted as a dominant-negative repressor, suggesting that Pbp is a genuine transcriptional co-activator for PPAR. The predicted 1,560-amino acid Pbp protein contains 2 LXXLL motifs, which are considered necessary and sufficient for the binding of several co-activators to nuclear receptors. Northern blot analysis detected Pbp expression in all mouse tissues examined, with higher levels in liver, kidney, lung, and testis. In situ hybridization showed that Pbp is expressed during mouse ontogeny, suggesting a possible role for Pbp in cellular proliferation and differentiation. In adult mouse, in situ hybridization detected Pbp expression in liver, bronchial epithelium in the lung, intestinal mucosa, kidney cortex, thymic cortex, splenic follicles, and seminiferous epithelium in testis. Lateron PPARBP was identified, which was called

TRAP220, from an immunopurified TR-alpha (THRA)-TRAP complex [Yuan et al., 1998, (11)]. The authors cloned Jurkat cell cDNAs encoding TRAP220. The predicted 1,581-amino acid TRAP220 protein contains LXXLL domains, which are found in other nuclear receptor-interacting proteins. TRAP220 is nearly identical to RB18A , with these proteins differing primarily by an extended N terminus on TRAP220. In the absence of TR-alpha, TRAP220 appears to reside in a single complex with other TRAPs. TRAP220 showed a direct ligand-dependent interaction with TR-alpha, which was mediated through the C terminus of TR-alpha and, at least in part, the LXXLL domains of TRAP220. TRAP220 also interacted with other nuclear receptors, including vitamin D receptor, RARA, RXRA, PPARA, PPARG, and estrogen receptor-alpha (ESR1; 133430), in a ligand-dependent manner. TRAP220 moderately stimulated human TR-alpha-mediated transcription in transfected cells, whereas a fragment containing the LXXLL motifs acted as a dominant-negative inhibitor of nuclear receptor-mediated transcription both in transfected cells and in cell-free transcription systems. Further studies indicated that TRAP220 plays a major role in anchoring other TRAPs to TR-alpha during the function of the TR-alpha-TRAP complex and that TRAP220 may be a global co-activator for the nuclear receptor superfamily. PBP, a nuclear receptor co-activator, interacts with estrogen receptor-alpha (ESR1) in the absence of estrogen. This interaction was enhanced in the presence of estrogen, but was reduced in the presence of the anti-estrogen Tamoxifen. Transfection of PBP into cultured cells resulted in enhancement of estrogen-dependent transcription, indicating that PBP serves as a co-activator in estrogen receptor signaling. To examine whether overexpression of PBP plays a role in breast cancer because of its co-activator function in estrogen receptor signaling, the levels of PBP expression in breast tumors was determined [Zhu et al., 1999, (12)]. High levels of PBP expression were detected in approximately 50% of primary breast cancers and breast cancer cell lines by ribonuclease protection analysis, in situ hybridization, and immunoperoxidase staining. By using FISH, the authors mapped the PBP gene to 17q12, a region that is amplified in some breast cancers. They found PBP gene amplification in approximately 24% (6 of 25) of breast tumors and approximately 30% (2 of 6) of breast cancer cell lines, implying that PBP gene overexpression can occur independent of gene amplification. They determined that the PBP gene comprises 17 exons that together span more than 37 kb. Their findings, in particular PBP gene amplification, suggested that PBP, by its ability to function as an estrogen receptor-alpha co-activator, may play a role in mammary epithelial differentiation and in breast carcinogenesis.

NEUROD2

Basic helix-loop-helix (bHLH) proteins are transcription factors involved in determining cell type during development. In 1995 a bHLH protein was described, termed NeuroD (for 'neurogenic

differentiation'), that functions during neurogenesis [Lee et al., 1995, (13)]. The human NEUROD gene maps to chromosome 2q32. The cloning and characterization of 2 additional NEUROD genes, NEUROD2 and NEUROD3 was described in 1996 [McCormick et al., 1996, (14)]. Sequences for the mouse and human homologues were presented. NEUROD2 shows a high degree of homology to the bHLH region of NEUROD, whereas NEUROD3 is more distantly related. The authors found that mouse neuroD2 was initially expressed at embryonic day 11, with persistent expression in the adult nervous system. Similar to neuroD, neuroD2 appears to mediate neuronal differentiation. The human NEUROD2 was mapped to 17q12 by fluorescence in situ hybridization and the mouse homologue to chromosome 11 [Tamimi et al., 1997, (15)].

TELETHONIN

Telethonin is a sarcomeric protein of 19 kD found exclusively in striated and cardiac muscle It appears to be localized to the Z disc of adult skeletal muscle and cultured myocytes. Telethonin is a substrate of titin, which acts as a molecular 'ruler' for the assembly of the sarcomere by providing spatially defined binding sites for other sarcomeric proteins. After activation by phosphorylation and calcium/calmodulin binding, titin phosphorylates the C-terminal domain of telethonin in early differentiating myocytes. The telethonin gene has been mapped to 17q12, adjacent to the phenylethanolamine N-methyltransferase gene [Valle et al., 1997, (16)].

PENT, PNMT

Phenylethanolamine N-methyltransferase catalyzes the synthesis of epinephrine from norepinephrine, the last step of catecholamine biosynthesis. The cDNA clone was first isolated in 1998 for bovine adrenal medulla PNMT using mixed oligodeoxyribonucleotide probes whose synthesis was based on the partial amino acid sequence of tryptic peptides from the bovine enzyme [Kaneda et al., 1988, (17)]. Using a bovine cDNA as a probe, the authors screened a human pheochromocytoma cDNA library and isolated a cDNA clone with an insert of about 1.0 kb, which contained a complete coding region of the enzyme. Northern blot analysis of human pheochromocytoma polyadenylated RNA using this cDNA insert as the probe demonstrated a single RNA species of about 1,000 nucleotides, suggesting that this clone is a full-length cDNA. The nucleotide sequence showed that human PNMT has 282 amino acid residues with a predicted molecular weight of 30,853, including the initial methionine. The amino acid sequence was 88% homologous to that of bovine enzyme. The PNMT gene was found to consist of 3 exons and 2 introns spanning about 2,100 basepairs. It was demonstrated that in transgenic mice the gene is expressed in adrenal medulla and retina. A hybrid gene consisting of 2 kb of the PNMT 5-prime-flanking region fused to the simian virus 40 early region also resulted in tumor antigen mRNA

expression in adrenal glands and eyes; furthermore, immunocytochemistry showed that the tu antigen was localized in nuclei of adrenal medullary cells and cells of the inner nuclear cell l of the retina, both prominent sites of epinephrine synthesis. The results indicate that enhancer(s) for appropriate expression of the gene in these cell types are in the 2-kb 5-pri flanking region of the gene.

Kaneda et al., 1988 (17), assigned the human PNMT gene to chromosome 17 by Southern analysis of DNA from mouse-human somatic cell hybrids. In 1992 the localization was narro down to 17q21-q22 by linkage analysis using RFLPs related to the PNMT gene and several DNA markers [Hoehe et al., 1992, (18)]. The findings are of interest in light of the descriptior a genetic locus associated with blood pressure regulation in the stroke-prone spontaneo hypertensive rat (SHR-SP) on rat chromosome 10 in a conserved linkage synteny group cc sponding to human chromosome 17q22-q24. See essential hypertension .

MGC9753

This gene maps on chromosome 17, at 17q12 according to RefSeq. It is expressed at very l level. It is defined by cDNA clones and produces, by alternative splicing, 7 different transc can be obtained (SEQ ID NO:60 to 66 and 83 to 89 ,Table 1), altogether encoding 7 diffe protein isoforms. Of specific interest is the putatively secreted isoform g, encoded by a mRN 2.55 kb. It's premessenger covers 16.94 kb on the genome. It has a very long 3' UTR. . The pro (226 aa, MW 24.6 kDa, pI 8.5) contains no Pfam motif. The MGC9753 gene produces, alternative splicing, 7 types of transcripts, predicted to encode 7 distinct proteins. It contain confirmed introns, 10 of which are alternative. Comparison to the genome sequence shows tha introns follow the consensual [gt-ag] rule, 1 is atypical with good support [tg_cg]. The six n abundant isoforms are designated by a) to i) and code for proteins as follows:

a) This mRNA is 3.03 kb long, its premessenger covers 16.95 kb on the genome. It has a long 3' UTR. The protein (190 aa, MW 21.5 kDa, pI 7.2) contains no Pfam motif. predicted to localise in the endoplasmic reticulum.

c) This mRNA is 1.17 kb long, its premessenger covers 16.93 kb on the genome. It ma incomplete at the N terminus. The protein (368 aa, MW 41.5 kDa, pI 7.3) contain Pfam motif.

d) This mRNA is 3.17 kb long, its premessenger covers 16.94 kb on the genome. It has a long 3' UTR and 5'p UTR. . The protein (190 aa, MW 21.5 kDa, pI 7.2) contains no P motif. It is predicted to localise in the endoplasmic reticulum.

g) This mRNA is 2.55 kb long, its premessenger covers 16.94 kb on the genome. It has a very long 3' UTR. . The protein (226 aa, MW 24.6 kDa, pI 8.5) contains no Pfam motif. It is predicted to be secreted.

h) This mRNA is 2.68 kb long, its premessenger covers 16.94 kb on the genome. It has a very long 3' UTR. . The protein (320 aa, MW 36.5 kDa, pI 6.8) contains no Pfam motif. It is predicted to localise in the endoplasmic reticulum.

i) This mRNA is 2.34 kb long, its premessenger covers 16.94 kb on the genome. It may be incomplete at the N terminus. It has a very long 3' UTR. . The protein (217 aa, MW 24.4 kDa, pI 5.9) contains no Pfam motif.

The MCG9753 gene may be homologue to the CAB2 gene located on chromosome 17q12. The CAB2, a human homologue of the yeast COS16 required for the repair of DNA double-strand breaks was cloned. Autofluorescence analysis of cells transfected with its GFP fusion protein demonstrated that CAB2 translocates into vesicles, suggesting that overexpression of CAB2 may decrease intercellular Mn-

(2 +) by accumulating it in the vesicles, in the same way as yeast.

Her-2/neu, ERBB2, NGL, TKR1

The oncogene originally called NEU was derived from rat neuro/glioblastoma cell lines. It encodes a tumor antigen, p185, which is serologically related to EGFR, the epidermal growth factor receptor. EGFR maps to chromosome 7. In1985 it was found, that the human homologue, which they designated NGL (to avoid confusion with neuraminidase, which is also symbolized NEU), maps to 17q12-q22 by in situ hybridization and to 17q21-qter in somatic cell hybrids [Yang-Feng et al., 1985, (19)]. Thus, the SRO is 17q21-q22. Moreover, in1985 a potential cell surface receptor of the tyrosine kinase gene family was identified and characterized by cloning the gene [Coussens et al., 1985, (20)]. Its primary sequence is very similar to that of the human epidermal growth factor receptor. Because of the seemingly close relationship to the human EGF receptor, the authors called the gene HER2. By Southern blot analysis of somatic cell hybrid DNA and by in situ hybridization, the gene was assigned to 17q21-q22. This chromosomal location of the gene is coincident with the NEU oncogene, which suggests that the 2 genes may in fact be the same; indeed, sequencing indicates that they are identical. In1988 a correlation between overexpression of NEU protein and the large-cell, comedo growth type of ductal carcinoma was found [van de Vijver et al., 1988, (21)]. The authors found no correlation, however, with lymph-node status or tumor recurrence. The role of HER2/NEU in breast and ovarian cancer was described in 1989,

which together account for one-third of all cancers in women and approximately one-quarter cancer-related deaths in females [Slamon et al., 1989, (22)].

An ERBB-related gene that is distinct from the ERBB gene, called ERBB1 was found in 19 ERBB2 was not amplified in vulva carcinoma cells with EGFR amplification and did not re with EGF receptor mRNA. About 30-fold amplification of ERBB2 was observed in a hum adenocarcinoma of the salivary gland. By chromosome sorting combined with veloc sedimentation and Southern hybridization, the ERBB2 gene was assigned to chromosome [Fukushige et al.,1986, (23)]. By hybridization to sorted chromosomes and to metaphase sprea with a genomic probe, they mapped the ERBB2 locus to 17q21. This is the chromosome breakpoint in acute promyelocytic leukemia (APL). Furthermore, they observed amplification ar elevated expression of the ERBB2 gene in a gastric cancer cell line. Antibodies against a synthet peptide corresponding to 14 amino acid residues at the COOH-terminus of a protein deduced fro the ERBB2 nucleotide sequence were raised in 1986. With these antibodies, the ERBB2 ger product from adenocarcinoma cells was precipitated and demonstrated to be a 185-k glycoprotein with tyrosine kinase activity. A cDNA probe for ERBB2 and by in situ hybridizatic to APL cells with a 15;17 chromosome translocation located the gene to the proximal side of th breakpoint [Kaneko et al., 1987, (24)]. The authors suggested that both the gene and the breakpoir are located in band 17q21.1 and, further, that the ERBB2 gene is involved in the development c leukemia. In 1987 experiments indicated that NEU and HER2 are both the same as ERBB2 [D Fiore et al., 1987, (25)]. The authors demonstrated that overexpression alone can convert the gen for a normal growth factor receptor, namely, ERBB2, into an oncogene. The ERBB2 to 17q11-q2 by in situ hybridization [Popescu et al., 1989, (26)]. By in situ hybridization to chromosome derived from fibroblasts carrying a constitutional translocation between 15 and 17, they showe that the ERBB2 gene was relocated to the derivative chromosome 15; the gene can thus be localized to 17q12-q21.32. By family linkage studies using multiple DNA markers in the 17q12-q21 region the ERBB2 gene was placed on the genetic map of the region.

Interleukin-6 is a cytokine that was initially recognized as a regulator of immune and inflammatory responses, but also regulates the growth of many tumor cells, including prostate cancer. Overexpression of ERBB2 and ERBB3 has been implicated in the neoplastic transformation of prostate cancer. Treatment of a prostate cancer cell line with IL6 induced tyrosine phosphorylation of ERBB2 and ERBB3, but not ERBB1/EGFR. The ERBB2 forms a complex with the gp130 subunit of the IL6 receptor in an IL6-dependent manner. This association was important because the inhibition of ERBB2 activity resulted in abrogation of IL6-induced MAPK activation. Thus, ERBB2 is a critical component of IL6 signaling through the MAP kinase

pathway [Qiu et al., 1998, (27)]. These findings showed how a cytokine receptor can diversify its signaling pathways by engaging with a growth factor receptor kinase.

Overexpression of ERBB2 confers Taxol resistance in breast cancers. Overexpression of ERBB2 inhibits Taxol-induced apoptosis [Yu et al., 1998, (28)]. Taxol activates CDC2 kinase in MDA-MB-435 breast cancer cells, leading to cell cycle arrest at the G2/M phase and, subsequently, apoptosis. A chemical inhibitor of CDC2 and a dominant-negative mutant of CDC2 blocked Taxol-induced apoptosis in these cells. Overexpression of ERBB2 in MDA-MB-435 cells by transfection transcriptionally upregulates CDKN1A which associates with CDC2, inhibits Taxol-mediated CDC2 activation, delays cell entrance to G2/M phase, and thereby inhibits Taxol-induced apoptosis. In CDKN1A antisense-transfected MDA-MB-435 cells or in p21-/- MEF cells, ERBB2 was unable to inhibit Taxol-induced apoptosis. Therefore, CDKN1A participates in the regulation of a G2/M checkpoint that contributes to resistance to Taxol-induced apoptosis in ERBB2-overexpressing breast cancer cells.

A secreted protein of approximately 68 kD was described, designated herstatin, as the product of an alternative ERBB2 transcript that retains intron 8 [Doherty et al., 1999, (29)]. This alternative transcript specifies 340 residues identical to subdomains I and II from the extracellular domain of p185ERBB2, followed by a unique C-terminal sequence of 79 amino acids encoded by intron 8. The recombinant product of the alternative transcript specifically bound to ERBB2-transfected cells and was chemically crosslinked to p185ERBB2, whereas the intron-encoded sequence alone also bound with high affinity to transfected cells and associated with p185 solubilized from cell extracts. The herstatin mRNA was expressed in normal human fetal kidney and liver, but was at reduced levels relative to p185ERBB2 mRNA in carcinoma cells that contained an amplified ERBB2 gene. Herstatin appears to be an inhibitor of p185ERBB2, because it disrupts dimers, reduces tyrosine phosphorylation of p185, and inhibits the anchorage-independent growth of transformed cells that overexpress ERBB2. The HER2 gene is amplified and HER2 is overexpressed in 25 to 30% of breast cancers, increasing the aggressiveness of the tumor. Finally, it was found that a recombinant monoclonal antibody against HER2 increased the clinical benefit of first-line chemotherapy in metastatic breast cancer that overexpresses HER2 [Slamon et al., 2001, (30)].

GRB7

Growth factor receptor tyrosine kinases (GF-RTKs) are involved in activating the cell cycle. Several substrates of GF-RTKs contain Src-homology 2 (SH2) and SH3 domains. SH2 domain-containing proteins are a diverse group of molecules important in tyrosine kinase signaling. Using

the CORT (cloning of receptor targets) method to screen a high expression mouse library, the gene for murine Grb7, which encodes a protein of 535 amino acids, was isolated [Margolis et al., 1992, (31)]. GRB7 is homologous to ras-GAP (ras-GTPase-activating protein). It contains an SH2 domain and is highly expressed in liver and kidney. This gene defines the GRB7 family, whose members include the mouse gene Grb10 and the human gene GRB14.

A putative GRB7 signal transduction molecule and a GRB7V novel splice variant from an invasive human esophageal carcinoma was isolated [Tanaka et al., 1998, (32)]. Although both GRB7 isoforms shared homology with the Mig-10 cell migration gene of Caenorhabditis elegans, the GRB7V isoform lacked 88 basepairs in the C terminus; the resultant frameshift led to substitution of an SH2 domain with a short hydrophobic sequence. The wildtype GRB7 protein, but not the GRB7V isoform, was rapidly tyrosyl phosphorylated in response to EGF stimulation in esophageal carcinoma cells. Analysis of human esophageal tumor tissues and regional lymph nodes with metastases revealed that GRB7V was expressed in 40% of GRB7-positive esophageal carcinomas. GRB7V expression was enhanced after metastatic spread to lymph nodes as compared to the original tumor tissues. Transfection of an antisense GRB7 RNA expression construct lowered endogenous GRB7 protein levels and suppressed the invasive phenotype exhibited by esophageal carcinoma cells. These findings suggested that GRB7 isoforms are involved in cell invasion and metastatic progression of human esophageal carcinomas. By sequence analysis, The GRB7 gene was mapped to chromosome 17q21-q22, near the topoisomerase-2 gene [Dong et al., 1997, (33)]. GRB-7 is amplified in concert with HER2 in several breast cancer cell lines and that GRB-7 is overexpressed in both cell lines and breast tumors. GRB-7, through its SH2 domain, binds tightly to HER2 such that a large fraction of the tyrosine phosphorylated HER2 in SKBR-3 cells is bound to GRB-7 [Stein et al., 1994, (34)].

GCSF, CSF3

Granulocyte colony-stimulating factor (or colony stimulating factor-3) specifically stimulates the proliferation and differentiation of the progenitor cells for granulocytes. The partial amino acid sequence of purified GCSF protein was determined, and by using oligonucleotides as probes, several GCSF cDNA clones were isolated from a human squamous carcinoma cell line cDNA library [Nagata et al., 1986, (35)]. Cloning of human GCSF cDNA shows that a single gene codes for a 177- or 180-amino acid mature protein of molecular weight 19,600. The authors found that the GCSF gene has 4 introns and that 2 different polypeptides are synthesized from the same gene by differential splicing of mRNA. The 2 polypeptides differ by the presence or absence of 3 amino acids. Expression studies indicate that both have authentic GCSF activity. A stimulatory activity from a glioblastoma multiform cell line being biologically and biochemically indistinguishable

from GCSF produced by a bladder cell line was found in 1987. By somatic cell hybridization and in situ chromosomal hybridization, the GCSF gene was mapped to 17q11 in the region of the breakpoint in the 15;17 translocation characteristic of acute promyelocytic leukemia [Le Beau et al., 1987, (36)]. Further studies indicated that the gene is proximal to the said breakpoint and that it remains on the rearranged chromosome 17. Southern blot analysis using both conventional and pulsed field gel electrophoresis showed no rearranged restriction fragments. By use of a full-length cDNA clone as a hybridization probe in human-mouse somatic cell hybrids and in flow-sorted human chromosomes, the gene for GCSF was mapped to 17q21-q22 lateron

THRA, THRA1, ERBA, EAR7, ERBA2, ERBA3

Both human and mouse DNA have been demonstrated to have two distantly related classes of ERBA genes and that in the human genome multiple copies of one of the classes exist [Jansson et al., 1983, (37)]. A cDNA was isolated derived from rat brain messenger RNA on the basis of homology to the human thyroid receptor gene [Thompson et al., 1987, (38)]. Expression of this cDNA produced a high-affinity binding protein for thyroid hormones. Messenger RNA from this gene was expressed in tissue-specific fashion, with highest levels in the central nervous system and no expression in the liver. An increasing body of evidence indicated the presence of multiple thyroid hormone receptors. The authors suggested that there may be as many as 5 different but related loci. Many of the clinical and physiologic studies suggested the existence of multiple receptors. For example, patients had been identified with familial thyroid hormone resistance in which peripheral response to thyroid hormones is lost or diminished while neuronal functions are maintained. Thyroidologists recognize a form of cretinism in which the nervous system is severely affected and another form in which the peripheral functions of thyroid hormone are more dramatically affected.

The cDNA encoding a specific form of thyroid hormone receptor expressed in human liver, kidney, placenta, and brain was isolated [Nakai et al., 1988, (39)]. Identical clones were found in human placenta. The cDNA encodes a protein of 490 amino acids and molecular mass of 54,824. Designated thyroid hormone receptor type alpha-2 (THRA2), this protein is represented by mRNAs of different size in liver and kidney, which may represent tissue-specific processing of the primary transcript.

The THRA gene contains 10 exons spanning 27 kb of DNA. The last 2 exons of the gene are alternatively spliced. A 5-kb THRA1 mRNA encodes a predicted 410-amino acid protein; a 2.7-kb THRA2 mRNA encodes a 490-amino acid protein. A third isoform, TR-alpha-3, is derived by alternative splicing. The proximal 39 amino acids of the TH-alpha-2 specific sequences are deleted

in TR-alpha-3. A second gene, THRB on chromosome 3, encodes 2 isoforms of TR-beta l alternative splicing. In1989the structure and function of the EAR1 and EAR7 genes w elucidated, both located on 17q21 [Miyajima et al., 1989, (40)]. The authors determined that o of the exons in the EAR7 coding sequence overlaps an exon of EAR1, and that the 2 genes a transcribed from opposite DNA strands. In addition, the EAR7 mRNA generates 2 alternative spliced isoforms, referred to as EAR71 and EAR72, of which the EAR71 protein is the huma counterpart of the chicken c-erbA protein.

The thyroid hormone receptors, beta, alpha-1, and alpha-2 3 mRNAs are expressed in all tissue examined and the relative amounts of the three mRNAs were roughly parallel. None of the mRNAs was abundant in liver, which is the major thyroid hormone-responsive organ. This led t the assumption that another thyroid hormone receptor may be present in liver. It was found tha ERBA, which potentiates ERBB, has an amino acid sequence different from that of other know oncogene products and related to those of the carbonic anhydrases [Debuire et al., 1984, (41)] ERBA potentiates ERBB by blocking differentiation of erythroblasts at an immature stage Carbonic anhydrases participate in the transport of carbon dioxide in erythrocytes. In 1986 it wa shown that the ERBA protein is a high-affinity receptor for thyroid hormone. The cDNA sequence indicates a relationship to steroid-hormone receptors, and binding studies indicate that it is a receptor for thyroid hormones. It is located in the nucleus, where it binds to DNA and activates transcription.

Maternal thyroid hormone is transferred to the fetus early in pregnancy and is postulated to regulate brain development. The ontogeny of TR isoforms and related splice variants in 9 first-trimester fetal brains by semi-quantitative RT-PCR analysis has been investigated. Expression of the TR-beta-1, TR-alpha-1, and TR-alpha-2 isoforms was detected from 8.1 weeks' gestation. An additional truncated species was detected with the TR-alpha-2 primer set, consistent with the TR-alpha-3 splice variant described in the rat. All TR-alpha-derived transcripts were coordinately expressed and increased approximately 8-fold between 8.1 and 13.9 weeks' gestation. A more complex ontogenic pattern was observed for TR-beta-1, suggestive of a nadir between 8.4 and 12.0 weeks' gestation. The authors concluded that these findings point to an important role for the TR-alpha-1 isoform in mediating maternal thyroid hormone action during first-trimester fetal brain development.

The identification of the several types of thyroid hormone receptor may explain the normal variation in thyroid hormone responsiveness of various organs and the selective tissue abnormalities found in the thyroid hormone resistance syndromes. Members of sibships, who were resistant to thyroid hormone action, had retarded growth, congenital deafness, and abnormal

bones, but had normal intellect and sexual maturation, as well as augmented cardiovascular activity. In this family abnormal T3 nuclear receptors in blood cells and fibroblasts have been demonstrated. The availability of cDNAs encoding the various thyroid hormone receptors was considered useful in determining the underlying genetic defect in this family.

The ERBA oncogene has been assigned to chromosome 17. The ERBA locus remains on chromosome 17 in the t(15;17) translocation of acute promyelocytic leukemia (APL). The thymidine kinase locus is probably translocated to chromosome 15; study of leukemia with t(17;21) and apparently identical breakpoint showed that TK was on 21q+. By in situ hybridization of a cloned DNA probe of c-erb-A to meiotic pachytene spreads obtained from uncultured spermatocytes it has been concluded that ERBA is situated at 17q21.33-17q22, in the same region as the break that generated the t(15;17) seen in APL. Because most of the grains were seen in 17q22, they suggested that ERBA is probably in the proximal region of 17q22 or at the junction between 17q22 and 17q21.33. By in situ hybridization it has been demonstrated, that that ERBA remains at 17q11-q12 in APL, whereas TP53, at 17q21-q22, is translocated to chromosome 15. Thus, ERBA must be at 17q11.2 just proximal to the breakpoint in the APL translocation and just distal to it in the constitutional translocation.

The aberrant THRA expression in nonfunctioning pituitary tumors has been hypothesized to reflect mutations in the receptor coding and regulatory sequences. They screened THRA mRNA and THRB response elements and ligand-binding domains for sequence anomalies. Screening THRA mRNA from 23 tumors by RNAse mismatch and sequencing candidate fragments identified 1 silent and 3 missense mutations, 2 in the common THRA region and 1 that was specific for the alpha-2 isoform. No THRB response element differences were detected in 14 nonfunctioning tumors, and no THRB ligand-binding domain differences were detected in
nonfunctioning tumors. Therefore it has been suggested that the novel thyroid receptor mutations may be of functional significance in terms of thyroid receptor action, and further definition of their functional properties may provide insight into the role of thyroid receptors in growth control in pituitary cells.

RAR-alpha

A cDNA encoding a protein that binds retinoic acid with high affinity has been cloned [Petkovich et al., 1987, (42)]. The protein was found to be homologous to the receptors for steroid hormones, thyroid hormones, and vitamin D3, and appeared to be a retinoic acid-inducible transacting enhancer factor. Thus, the molecular mechanisms of the effect of vitamin A on embryonic development, differentiation and tumor cell growth may be similar to those described for other

members of this nuclear receptor family. In general, the DNA-binding domain is most high conserved, both within and between the 2 groups of receptors (steroid and thyroid); Using a cDN probe, the RAR-alpha gene has been mapped to 17q21 by in situ hybridization [Mattei et al., 198 (43)]. Evidence has been presented for the existence of 2 retinoic acid receptors, RAR-alpha ar RAR-beta, mapping to chromosome 17q21.1 and 3p24, respectively. The alpha and beta forms (RAR were found to be more homologous to the 2 closely related thyroid hormone receptors alph and beta, located on 17q11.2 and 3p25-p21, respectively, than to any other members of the nucle receptor family. These observations suggest that the thyroid hormone and retinoic acid receptoı evolved by gene, and possibly chromosome, duplications from a common ancestor, which itsel diverged rather early in evolution from the common ancestor of the steroid receptor group of th family. They noted that the counterparts of the human RARA and RARB genes are present in bot the mouse and chicken. The involvement of RARA at the APL breakpoint may explain why th use of retinoic acid as a therapeutic differentiation agent in the treatment of acute myeloi leukemias is limited to APL. Almost all patients with APL have a chromosomal translocatio t(15;17)(q22;q21). Molecular studies reveal that the translocation results in a chimeric gen through fusion between the PML gene on chromosome 15 and the RARA gene on chromosom 17. A hormone-dependent interaction of the nuclear receptors RARA and RXRA with CLOCK and MOP4 has been presented.

CDC18 L, CDC 6

In yeasts, Cdc6 (Saccharomyces cerevisiae) and Cdc18 (Schizosaccharomyces pombe) associate with the origin recognition complex (ORC) proteins to render cells competent for DNA replication. Thus, Cdc6 has a critical regulatory role in the initiation of DNA replication in yeast. cDNAs encoding Xenopus and human homologues of yeast CDC6 have been isolated [Williams et al., 1997, (44)]. They designated the human and Xenopus proteins p62(cdc6). Independently, in a yeast 2-hybrid assay using PCNA as bait, cDNAs encoding the human CDC6/Cdc18 homologue have been isolated [Saha et al, 1998, (45)]. These authors reported that the predicted 560-amino acid human protein shares approximately 33% sequence identity with the 2 yeast proteins. On Western blots of HeLa cell extracts, human CDC6/cdc18 migrates as a 66-kD protein. Although Northern blots indicated that CDC6/Cdc18 mRNA levels peak at the onset of S phase and diminish at the onset of mitosis in HeLa cells, the authors found that total CDC6/Cdc18 protein level is unchanged throughout the cell cycle. Immunofluorescent analysis of epitope-tagged protein revealed that human CDC6/Cdc18 is nuclear in G1- and cytoplasmic in S-phase cells, suggesting that DNA replication may be regulated by either the translocation of this protein between the nucleus and cytoplasm or by selective degradation of the protein in the nucleus.

Immunoprecipitation studies showed that human CDC6/Cdc18 associates in vivo with cyclin A, CDK2,and ORC1. The association of cyclin-CDK2 with CDC6/Cdc18 was specifically inhibited by a factor present in mitotic cell extracts. Therefore it has been suggested that if the interaction between CDC6/Cdc18 with the S phase-promoting factor cyclin-CDK2 is essential for the initiation of DNA replication, the mitotic inhibitor of this interaction could prevent a premature interaction until the appropriate time in G1. Cdc6 is expressed selectively in proliferating but not quiescent mammalian cells, both in culture and within tissues in intact animals [Yan et al., 1998, (46)]. During the transition from a growth-arrested to a proliferative state, transcription of mammalian Cdc6 is regulated by E2F proteins, as revealed by a functional analysis of the human Cdc6 promoter and by the ability of exogenously expressed E2F proteins to stimulate the endogenous Cdc6 gene. Immunodepletion of Cdc6 by microinjection of anti-Cdc6 antibody blocked initiation of DNA replication in a human tumor cell line. The authors concluded that expression of human Cdc6 is regulated in response to mitogenic signals through transcriptional control mechanisms involving E2F proteins, and that Cdc6 is required for initiation of DNA replication in mammalian cells.

Using a yeast 2-hybrid system, co-purification of recombinant proteins, and immunoprecipitation, it has been demonstrated lateron that an N-terminal segment of CDC6 binds specifically to PR48, a regulatory subunit of protein phosphatase 2A (PP2A). The authors hypothesized that dephosphorylation of CDC6 by PP2A, mediated by a specific interaction with PR48 or a related B-double prime protein, is a regulatory event controlling initiation of DNA replication in mammalian cells. By analysis of somatic cell hybrids and by fluorescence in situ hybridization the human p62(cdc6) gene has been to 17q21.3.

TOP2A, TOP2, TOP2 alpha

DNA topoisomerases are enzymes that control and alter the topologic states of DNA in both prokaryotes and eukaryotes. Topoisomerase II from eukaryotic cells catalyzes the relaxation of supercoiled DNA molecules, catenation, decatenation, knotting, and unknotting of circular DNA. It appears likely that the reaction catalyzed by topoisomerase II involves the crossing-over of 2 DNA segments. It has been estimated that there are about 100,000 molecules of topoisomerase II per HeLa cell nucleus, constituting about 0.1% of the nuclear extract. Since several of the abnormal characteristics of ataxia-telangiectasia appear to be due to defects in DNA processing, screening for these enzyme activities in 5 AT cell lines has been performed [Singh et al., 1988, (47)]. In comparison to controls, the level of DNA topoisomerase II, determined by unknotting of P4 phage DNA, was reduced substantially in 4 of these cell lines and to a lesser extent in the fifth.

DNA topoisomerase I, assayed by relaxation of supercoil DNA, was found to be present at nor levels.

The entire coding sequence of the human TOP2 gene has been determined [Tsai-Pflugfelder et 1988, (48)].

In addition human cDNAs that had been isolated by screening a cDNA library derived fro: mechlorethamine-resistant Burkitt lymphoma cell line (Raji-HN2) with a Drosophila Top cDNA had been sequenced [Chung et al., 1989, (49)]. The authors identified 2 classes of seque representing 2 TOP2 isoenzymes, which have been named TOP2A and TOP2B. The sequence of the TOP2A cDNAs is identical to that of an internal fragment of the TOP2 cDNA isolatec Tsai-Pflugfelder et al., 1988 (48). Southern blot analysis indicated that the TOP2A and TO] cDNAs are derived from distinct genes. Northern blot analysis using a TOP2A-specific pr detected a 6.5-kb transcript in the human cell line U937. Antibodies against a TOP2A pep recognized a 170-kD protein in U937 cell lysates. Therefore it was concluded that their provide genetic and immunochemical evidence for 2 TOP2 isozymes. The complete structure: the TOP2A and TOP2B genes has been reported [Lang et al., 1998, (50)]. The TOP2A gene sp approximately 30 kb and contains 35 exons.

Tsai-Pflugfelder et al., 1988 (48) showed that the human enzyme is encoded by a single-copy g which they mapped to 17q21-q22 by a combination of in situ hybridization of a cloned fragmen metaphase chromosomes and by Southern hybridization analysis with a panel of mouse-hur hybrid cell lines. The assignment to chromosome 17 has been confirmed by the study of som cell hybrids. Because of co-amplification in an adenocarcinoma cell line, it was concluded that TOP2A and ERBB2 genes may be closely linked on chromosome 17 [Keith et al., 1992, (5 Using probes that detected RFLPs at both the TOP2A and TOP2B loci, the demonstr heterozygosity at a frequency of 0.17 and 0.37 for the alpha and beta loci, respectively. The mc homologue was mapped to chromosome 11 [Kingsmore et al., 1993, (52)]. The structure function of type II DNA topoisomerases has been reviewed [Watt et al., 1994, (53)]. D topoisomerase II-alpha is associated with the pol II holoenzyme and is a required componen chromatin-dependent co-activation. Specific inhibitors of topoisomerase II blocked transcrip on chromatin templates, but did not affect transcription on naked templates. Addition of puri topoisomerase II-alpha reconstituted chromatin-dependent activation activity in reactions core pol II. Therefore the transcription on chromatin templates seems to result in the accumula of superhelical tension, making the relaxation activity of topoisomerase II essential for produc RNA synthesis on nucleosomal DNA.

IGFBP4

Six structurally distinct insulin-like growth factor binding proteins have been isolated and their cDNAs cloned: IGFBP1, IGFBP2, IGFBP3, IGFBP4, IGFBP5 and IGFBP6. The proteins display strong sequence homologies, suggesting that they are encoded by a closely related family of genes. The IGFBPs contain 3 structurally distinct domains each comprising approximately one-third of the molecule. The N-terminal domain 1 and the C-terminal domain 3 of the 6 human IGFBPs show moderate to high levels of sequence identity including 12 and 6 invariant cysteine residues in domains 1 and 3, respectively (IGFBP6 contains 10 cysteine residues in domain 1), and are thought to be the IGF binding domains. Domain 2 is defined primarily by a lack of sequence identity among the 6 IGFBPs and by a lack of cysteine residues, though it does contain 2 cysteines in IGFBP4. Domain 3 is homologous to the thyroglobulin type I repeat unit. Recombinant human insulin-like growth factor binding proteins 4, 5, and 6 have been characterized by their expression in yeast as fusion proteins with ubiquitin [Kiefer et al., 1992, (54)]. Results of the study suggested to the authors that the primary effect of the 3 proteins is the attenuation of IGF activity and suggested that they contribute to the control of IGF-mediated cell growth and metabolism. Moreover, IGFBPs have influence on EGFR and Her-2/neu mediated signaling. Addition of IGFBPs to Her-2/neu overexpressing cells at least in part blocks growth and survival characteristica of the respective cells.

Based on peptide sequences of a purified insulin-like growth factor-binding protein (IGFBP) rat IGFBP4 has been cloned by using PCR [Shimasaki et al., 1990, (55)]. They used the rat cDNA to clone the human ortholog from a liver cDNA library. Human IGFBP4 encodes a 258-amino acid polypeptide, which includes a 21-amino acid signal sequence. The protein is very hydrophilic, which may facilitate its ability as a carrier protein for the IGFs in blood. Northern blot analysis of rat tissues revealed expression in all tissues examined, with highest expression in liver. It was stated that IGFBP4 acts as an inhibitor of IGF-induced bone cell proliferation. The genomic region containing the IGFBP gene. The gene consists of 4 exons spanning approximately 15 kb of genomic DNA has been examined [Zazzi et al., 1998, (56)]. The upstream region of the gene contains a TATA box and a cAMP-responsive promoter.

By in situ hybridization, the IGFBP4 gene was mapped to 17q12-q21 [Bajalica et al., 1992, (57)]. Because the hereditary breast-ovarian cancer gene BRCA1 had been mapped to the same region, it has been investigated whether IGFBP4 is a candidate gene by linkage analysis of 22 BRCA1 families; the finding of genetic recombination suggested that it is not the BRCA1 gene [Tonin et al., 1993, (58)].

EBI 1, CCR7, CMKBR7

Using PCR with degenerate oligonucleotides, a lymphoid-specific member of the G protei coupled receptor family has been identified and mapped mapped to 17q12-q21.2 by analysis human/mouse somatic cell hybrid DNAs and fluorescence in situ hybridization. It has been show that this receptor had been independently identified as the Epstein-Barr-induced cDNA (symb EBI1) [Birkenbach et al., 1993, (59)]. EBI1 is expressed in normal lymphoid tissues and in sever B- and T-lymphocyte cell lines. While the function and the ligand for EBI1 remains unknown, i sequence and gene structure suggest that it is related to receptors that recognize chemoattractant such as interleukin-8, RANTES, C5a, and fMet-Leu-Phe. Like the chemoattractant receptors, EBI contains intervening sequences near its 5-prime end; however, EBI1 is unique in that both of it introns interrupt the coding region of the first extracellular domain. Mouse Ebi1 cDNA has bee isolated and found to encode a protein with 86% identity to the human homologue.

Subsets of murine CD4+ T cells localize to different areas of the spleen after adoptive transfer Naive and T helper-1 (TH1) cells, which express CCR7, home to the periarteriolar lymphoi sheath, whereas activated TH2 cells, which lack CCR7, form rings at the periphery of the T-cel zones near B-cell follicles. It has been found that retroviral transduction of TH2 cells with CCR forced them to localize in a TH1-like pattern and inhibited their participation in B-cell help in viv but not in vitro. Apparently differential expression of chemokine receptors results in unique cellular migration patterns that are important for effective immune responses.

CCR7 expression divides human memory T cells into 2 functionally distinct subsets. CCR7- memory cells express receptors for migration to inflamed tissues and display immediate effector function. In contrast, $CCR7^+$ memory cells express lymph node homing receptors and lack immediate effector function, but efficiently stimulate dendritic cells and differentiate into $CCR7^-$ effector cells upon secondary stimulation. The $CCR7^+$ and $CCR7^-$ T cells, named central memory (T-CM) and effector memory (T-EM), differentiate in a step-wise fashion from naive T cells, persist for years after immunization, and allow a division of labor in the memory response.

CCR7 expression in memory $CD8^+$ T lymphocyte responses to HIV and to cytomegalovirus (CMV) tetramers has been evaluated. Most memory T lymphocytes express CD45RO, but a fraction express instead the CD45RA marker. Flow cytometric analyses of marker expression and cell division identified 4 subsets of HIV- and CMV-specific $CD8^+$ T cells, representing a lineage differentiation pattern: $CD45RA^+CCR7^+$ (double-positive); $CD45RA^-CCR7^+$; $CD45RA^-CCR7^-$ (double-negative); $CD45RA^+CCR7^-$. The capacity for cell division, as measured by 5-(and 6-)carboxyl-fluorescein diacetate, succinimidyl ester, and intracellular staining for the Ki67

nuclear antigen, is largely confined to the CCR7$^+$ subsets and occurred more rapidly in cells that are also CD45RA$^+$. Although the double-negative cells did not divide or expand after stimulation, they did revert to positivity for either CD45RA or CCR7 or both. The CD45RA$^+$CCR7$^-$ cells, considered to be terminally differentiated, fail to divide, but do produce interferon-gamma and express high levels of perforin. The representation of subsets specific for CMV and for HIV is distinct. Approximately 70% of HIV-specific CD8$^+$ memory T cells are double-negative or preterminally differentiated compared to 40% of CMV-specific cells. Approximately 50% of the CMV-specific CD8+ memory T cells are terminally differentiated compared to fewer than 10% of the HIV-specific cells. It has been proposed that terminally differentiated CMV-specific cells are poised to rapidly intervene, while double-positive precursor cells remain for expansion and replenishment of the effector cell pool. Furthermore, high-dose antigen tolerance and the depletion of HIV-specific CD4$^+$ helper T-cell activity may keep the HIV-specific memory CD8$^+$ T cells at the double-negative stage, unable to differentiate to the terminal effector state. B lymphocytes recirculate between B cell-rich compartments (follicles or B zones) in secondary lymphoid organs, surveying for antigen. After antigen binding, B cells move to the boundary of B and T zones to interact with T-helper cells. Furthermore it has been demonstrated that antigen-engaged B cells have increased expression of CCR7, the receptor for the T-zone chemokines CCL19 (also known as ELC) and CCL21, and that they exhibit increased responsiveness to both chemoattractants. In mice lacking lymphoid CCL19 and CCL21 chemokines, or with B cells that lack CCR7, antigen engagement fails to cause movement to the T zone. Using retroviral-mediated gene transfer, the authors demonstrated that increased expression of CCR7 is sufficient to direct B cells to the T zone. Reciprocally, overexpression of CXCR5, the receptor for the B-zone chemokine CXCL13, is sufficient to overcome antigen-induced B-cell movement to the T zone. This points toward a mechanism of B-cell relocalization in response to antigen, and established that cell position in vivo can be determined by the balance of responsiveness to chemoattractants made in separate but adjacent zones.

BAF57, SMARCE 1

The SWI/SNF complex in S. cerevisiae and Drosophila is thought to facilitate transcriptional activation of specific genes by antagonizing chromatin-mediated transcriptional repression. The complex contains an ATP-dependent nucleosome disruption activity that can lead to enhanced binding of transcription factors. The BRG1/brm-associated factors, or BAF, complex in mammals is functionally related to SWI/SNF and consists of 9 to 12 subunits, some of which are homologous to SWI/SNF subunits. A 57-kD BAF subunit, BAF57, is present in higher eukaryotes, but not in yeast. Partial coding sequence has been obtained from purified BAF57 from extracts of

a human cell line [Wang et al., 1998, (60)]. Based on the peptide sequences, they identified cDNAs encoding BAF57. The predicted 411-amino acid protein contains an HMG domain adjacent to a kinesin-like region. Both recombinant BAF57 and the whole BAF complex bind 4-way junction (4WJ) DNA, which is thought to mimic the topology of DNA as it enters or exits the nucleosome. The BAF57 DNA-binding activity has characteristics similar to those of other HMG proteins. It was found that complexes with mutations in the BAF57 HMG domain retain their DNA-binding and nucleosome-disruption activities. They suggested that the mechanism by which mammalian SWI/SNF-like complexes interact with chromatin may involve recognition of higher-order chromatin structure by 2 or more DNA-binding domains. RNase protection studies and Western blot analysis revealed that BAF57 is expressed ubiquitously. Several lines of evidence point toward the involvement of SWI/SNF factors in cancer development [Klochendler-Yeivin et al., 2002, (61)]. Moreover, SWI/SNF related genes are assigned to chromosomal regions that are frequently involved in somatic rearrangements in human cancers [Ring et al., 1998, (62)]. In this respect it is interesting that some of the SWI/SNF family members (i.e. SMARCC1, SMARCC2, SMARCD1 and SMARCD22 are neighboring 3 of the eucaryotic ARCHEONs we have identified (i.e. 3p21-p24, 12q13-q14 and 17q respectively)and which are part of the present invention. In this invention we could also map SMARCE1/BAF57 to the 17q12 region by PCR karyotyping.

KRT 10, K10

Keratin 10 is an intermediate filament (IF) chain which belongs to the acidic type I family and is expressed in terminally differentiated epidermal cells. Epithelial cells almost always co-express pairs of type I and type II keratins, and the pairs that are co-expressed are highly characteristic of a given epithelial tissue. For example, in human epidermis, 3 different pairs of keratins are expressed: keratins 5 (type II) and 14 (type I), characteristic of basal or proliferative cells; keratins 1 (type II) and 10 (type I), characteristic of superbasal terminally differentiating cells; and keratins 6 (type II) and 16 (type I) (and keratin 17 [type I]), characteristic of cells induced to hyper-proliferate by disease or injury, and epithelial cells grown in cell culture. The nucleotide sequence of a 1,700 bp cDNA encoding human epidermal keratin 10 (56.5 kD) [Darmon et al., 1987, (63)] has been published as well as the complete amino acid sequence of human keratin 10 [Zhou et al., 1988, (64)]. Polymorphism of the KRT10 gene, restricted to insertions and deletions of the glycine-richquasipeptide repeats that form the glycine-loop motif in the C-terminal domain, have been extensively described [Korge et al., 1992, (65)].

By use of specific cDNA clones in conjunction with somatic cell hybrid analysis and in situ hybridization, KRT10 gene has been mapped to 17q12-q21 in a region proximal to the breakpoint at 17q21 that is involved in a t(17;21)(q21;q22) translocation associated with a form of acute

leukemia. KRT10 appeared to be telomeric to 3 other loci that map in the same region: CSF3, ERBA1, and HER2 [Lessin et al., 1988, (66)]. NGFR and HOX2 are distal to K9. It has been demonstrated that the KRT10, KRT13, and KRT15 genes are located in the same large pulsed field gel electrophoresis fragment [Romano et al., 1991, (67)]. A correlation of assignments of the 3 genes makes 17q21-q22 the likely location of the cluster. Transgenic mice expressing a mutant keratin 10 gene have the phenotype of epidermolytic hyperkeratosis , thus suggesting that a genetic basis for the human disorder resides in mutations in genes encoding suprabasal keratins KRT1 or KRT10 [Fuchs et al 1992, (68)]. The authors also showed that stimulation of basal cell proliferation can result from a defect in suprabasal cells and that distortion of nuclear shape or alterations in cytokinesis can occur when an intermediate filament network is perturbed. In a family with keratosis palmaris et plantaris without blistering either spontaneously or in response to mild mechanical or thermal stress and with no involvement of the skin and parts of the body other than the palms and soles, a tight linkage to an insertion-deletion polymorphism in the C-terminal coding region of the KRT10 gene (maximum lod score = 8.36 at theta = 0.00) was found [Rogaev et al., 1993, (69)]. It is noteworthy that it was a rare, high molecular weight allele of the KRT10 polymorphism that segregated with the disorder. The allele was observed once in 96 independent chromosomes from unaffected Caucasians. The KRT10 polymorphism arose from the insertion/deletion of imperfect (CCG)n repeats within the coding region and gave rise to a variable glycine loop motif in the C-terminal tail of the keratin 10 protein. It is possible that there was a pathogenic role for the expansion of the imperfect trinucleotide repeat.

KRT12,K12

Keratins are a group of water-insoluble proteins that form 10 nm intermediate filaments in epithelial cells. Approximately 30 different keratin molecules have been identified. They can be divided into acidic and basic-neutral subfamilies according to their relative charges, immunoreactivity, and sequence homologies to types I and II wool keratins, respectively. In vivo, a basic keratin usually is co-expressed and 'paired' with a particular acidic keratin to form a heterodimer. The expression of various keratin pairs is tissue specific, differentiation dependent, and developmentally regulated. The presence of specific keratin pairs is essential for the maintenance of the integrity of epithelium. For example, mutations in human K14/K5 pair and the K10/K1 pair underlie the skin diseases, epidermolysis bullosa simplex and epidermolytic hyperkeratosis, respectively. Expression of the K3 and K12 keratin pair have been found in the cornea of a wide number of species, including human, mouse, and chicken, and is regarded as a marker for corneal-type epithelial differentiation. The murine Krt12 (Krt1.12) gene and demonstrated that its expression is corneal epithelial cell specific, differentiation dependent, and

developmentally regulated [Liu et al., 1993, (70)]. The corneal-specific nature of keratin 12 ge expression signifies keratin 12 plays a unique role in maintaining normal corneal epitheli function. Nevertheless, the exact function of keratin 12 remains unknown and no hereditary hum: corneal epithelial disorder has been linked directly to the mutation in the keratin 12 gene. As pa of a study of the expression profile of human corneal epithelial cells, a cDNA with an op reading frame highly homologous to the cornea-specific mouse keratin 12 gene has been isolat [Nishida et al., 1996, (71)]. To elucidate the function of keratin 12 knockout mice lacking th Krt1.12 gene have been created by gene targeting techniques. The heterozygous mice appeare normal. Homozygous mice developed normally and suffered mild corneal epithelial erosion. Th corneal epithelia were fragile and could be removed by gentle rubbing of the eyes or brushing. Th corneal epithelium of the homozygotes did not express keratin 12 as judged b immunohistochemistry, Western immunoblot analysis with epitope-specific anti-keratin 1. antibodies, Northern hybridization, and in situ hybridization with an antisense keratin 1: riboprobe. The KRT12 gene has been mapped to 17q by study of radiation hybrids and localized i to the type I keratin cluster in the interval between D17S800 and D17S930 (17q12-q21) [Nishid: et al., 1997, (72)]. The authors presented the exon-intron boundary structure of the KRT12 gen and mapped the gene to 17q12 by fluorescence in situ hybridization. The gene contains 7 introns defining 8 exons that cover the coding sequence. Together the exons and introns spar approximately 6 kb of genomic DNA.

Meesmann corneal dystrophy is an autosomal dominant disorder causing fragility of the anterior corneal epithelium, where the cornea-specific keratins K3 and K12 are expressed. Dominant-negative mutations in these keratins might be the cause of Meesmann corneal dystrophy. Indeed, linkage of the disorder to the K12 locus in Meesmann's original German kindred [Meesmann and Wilke, 1939, (73)] with Z(max) = 7.53 at theta = 0.0 has been found. In 2 pedigrees from Northern Ireland, they found that the disorder co-segregated with K12 in one pedigree and K3 in the other. Heterozygous missense mutations in K3 or in K12 (R135T, V143L,) in each family have been identified. All these mutations occurred in highly conserved keratin helix boundary motifs, where dominant mutations in other keratins have been found to compromise cytoskeletal function severely, leading to keratinocyte fragility.

The regions of the human KRT12 gene have been sequenced to enable mutation detection for all exons using genomic DNA as a template [Corden et al., 2000, (74)]. The authors found that the human genomic sequence spans 5,919 bp and consists of 8 exons. A microsatellite dinucleotide repeat was identified within intron 3, which was highly polymorphic and which they developed for use in genotype analysis. In addition, 2 mutations in the helix initiation motif of K12 were found

in families with Meesmann corneal dystrophy. In an American kindred, a missense M129T mutation was found in the KRT12 gene. They stated that a total of 8 mutations in the KRT12 gene had been reported.

Genetic interactions within ARCHEONs

Genes involved in genomic alterations (amplifications, insertions, translocations, deletions, etc.) exhibit changes in their expression pattern. Of particular interest are gene amplifications, which account for gene copy numbers >2 per cell or deletions accounting for gene copy numbers <2 per cell. Gene copy number and gene expression of the respective genes do not necessarily correlate. Transcriptional overexpression needs an intact transcriptional context, as determined by regulatory regions at the chromosomal locus (promotor, enhancer and silencer), and sufficient amounts of transcriptional regulators being present in effective combinations. This is especially true for genomic regions, which expression is tightly regulated in specific tissues or during specific developmental stages. ARCHEONs are specified by gene clusters of more than two genes being directly neighboured or in chromosomal order, interspersed by a maximum of 10, preferably 7, more preferably 5 or at least 1 gene. The interspersed genes are also co-amplified but do not directly interact with the ARCHEON. Such an ARCHEON may spread over a chromosomal region of a maximum of 20, more preferably 10 or at least 6 Megabases. The nature of an ARCHEON is characterized by the simultaneous amplification and/or deletion and the correlating expression (i.e. upregulation or downregulation respectively) of the encompassed genes in a specific tissue, cell type, cellular or developmental state or time point. Such ARCHEONs are commonly conserved during evolution, as they play critical roles during cellular development. In case of these ARCHEONs whole gene clusters are overexpressed upon amplification as they harbor self-regulatory feedback loops, which stabilize gene expression and/or biological effector functions even in abnormal biological settings, or are regulated by very similar transcription factor combinations, reflecting their simultaneous function in specific tissues at certain developmental stages. Therefore, the gene copy numbers correlates with the expression level especially for genes in gene clusters functioning as ARCHEONs. In case of abnormal gene expressions in neoplastic lesions it is of great importance to know whether the self-regulatory feedback loops have been conserved as they determine the biological activity of the ARCHEON gene members.

The intensive interaction between genes in ARCHEONs is described for the 17q21 ARCHEON (Fig. 1) by way of illustration not by limitation. In one embodiment the presence or absence of alterations of genes within distinct genomic regions are correlated with each other, as exemplified for breast cancer cell lines (Fig.3 and Fig. 4). This confers to the discovery of the present invention, that multiple interactions of said gene products of defined chromosomal localizations

happen, that according to their respective alterations in abnormal tissue have predictiv diagnostic, prognostic and/or preventive and therapeutic value. These interactions are mediate directly or indirectly, due to the fact that the respective genes are part of interconnected (independent signaling networks or regulate cellular behavior (differentiation status, proliferativ and /or apoptotic capacity, invasiveness, drug responsiveness, immune modulatory activities) in synergistic, antagonistic or independent fashion. The order of functionally important genes withi the ARCHEONs has been conserved during evolution (e.g. the ARCHEON on human chromosor 17q21 is present on mouse chromosome 11). Moreover, it has been found that the 17q2 ARCHEON is also present on human chromosome 3p21 and 12q13, both of which are als involved in amplification events and in tumor development. Most probably these homologou ARCHEONs were formed by duplications and rearrangements during vertebrate evolution Homologous ARCHEONs consist of homologous genes and/or isoforms of specific gene familie (e.g. RARA or RARB or RARG, THRA or THRB, TOP2A or TOP2B, RAB5A or RAB5B BAF170 or BAF 155, BAF60A or BAF60B, WNT5A or WNT5B, IGFBP4 or IGFBP6). Moreove these regions are flanked by homologous chromosomal gene clusters (e.g. CACN, SCYA, HOX Keratins). These ARCHEONs have diverged during evolution to fulfill their respective function in distinct tissues (e.g. the 17q21 ARCHEON has one of its main functions in the central nervous system). Due to their tissue specific function extensive regulatory loops control the expression of the members of each ARCHEON. During tumor development these regulations become critical for the characteristics of the abnormal tissues with respect to differentiation, proliferation, drug responsiveness, invasiveness. It has been found that the co-amplification of genes within ARCHEONs can lead to co-expression of the respective gene products. Some of said genes also exhibit additional mutations or specific patterns of polymorphisms, which are substantial for the oncogenic capacities of these ARCHEONs. It is one of the critical features of such amplicons, which members of the ARCHEON have been conserved during tumor formation (e.g. during amplification and deletion events), thereby defining these genes as diagnostic marker genes. Moreover, the expression of the certain genes within the ARCHEON can be influenced by other members of the ARCHEON, thereby defining the regulatory and regulated genes as target genes for therapeutic intervention. It was also observed, that the expression of certain members of the ARCHEON is sensitive to drug treatment (e.g. TOPO2 alpha, RARA, THRA, HER-2) which defines these genes as "marker genes". Moreover several other genes are suitable for therapeutic intervention by antibodies (CACNB1, EBI1), ligands (CACNB1) or drugs like e.g. kinase inhibitors (CrkRS, CDC6). The following examples of interactions between members of ARCHEONs are offered by way of illustration, not by way of limitation.

EBI1/CCR7 is lymphoid-specific member of the G protein-coupled receptor family. EBI1 recognizes chemoattractants, such as interleukin-8, SCYAs, Rantes, C5a, and fMet-Leu-Phe. The capacity for cell division is largely confined to the CCR7$^+$ subsets in lymphocytes. Double-negative cells did not divide or expand after stimulation. CCR7$^-$ cells, considered to be terminally differentiated, fail to divide, but do produce interferon-gamma and express high levels of perforin. EBI1 is induced by viral activities such as the Eppstein-Barr-Virus. Therefore, EBI1 is associated with transformation events in lymphocytes. A functional role of EBI1 during tumor formation in non-lymphoid tissues has been investigated in this invention. Interestingly, also ERBA and ERBB, located in the same genomic region, are associated with lymphocyte transformation. Moreover, ligands of the receptor (i.e. SCYA5/Rantes) are in genomic proximity on 17q. Abnormal expression of both of these factors in lymphoid and non-lymphoid tissues establishes an autorgulatory feedback loop, inducing signaling events within the respective cells. Expression of lymphoid factors has effect on immune cells and modulates cellular behavior. This is of particular interest with regard to abnormal breast tissue being infiltrated by lymphocytes. In line with this, another immunmodulatory and proliferation factor is located nearby on 17q21. Granulocyte colony-stimulating factor (GCSF3) specifically stimulates the proliferation and differentiation of the progenitor cells for granulocytes. A stimulatory activity from a glioblastoma multiforme cell line being biologically and biochemically indistinguishable from GCSF produced by a bladder cell line has also been found. Colony-stimulating factors not only affects immune cells, but also induce cellular responses of non-immune cells, indicating possible involvement in tumor development upon abnormal expression. In addition several other genes of the 17q21 ARCHEON are involved in proliferation, survival, differentiation of immune cells and/or lymphoblastic leukemia, such as MLLT6, ZNF144 and ZNFN1A3, again demonstrating the related functions of the gene products in interconnected key processes within specific cell types. Aberrant expression of more than one of these genes in non-immune cells constitutes signalling activities, that contribute to the oncogenic activities that derive solely from overexpression of the Her-2/neu gene.

PPARBP has been found in complex with the tumorsuppressor gene of the p53 family. Moreover, PPARBP also binds to PPAR-alpha (PPARA), RAR-alpha (RARA), RXR, THRA and TR-beta-1. Due to it's ability to bind to thyroid hormone receptors it has been named TRIP2 and TRAP220. In this complexes PPARBP affects gene regulatory activities. Interestingly, PPARBP is located in genomic proximity to its interaction partners THRA and RARA. We have found PPARBP to be co-amplified with THRA and RARA in tumor tissue. THRA has been isolated from avian erythroblastosis virus in conjunction with ERBB and therefore was named ERBA. ERBA potentiates ERBB by blocking differentiation of erythroblasts at an immature stage. ERBA has been shown to influence ERBB expression. In this setting deletions of C-terminal portions of the

THRA gene product are of influence. Aberrant THRA expression has also been found nonfunctioning pituitary tumors, which has been hypothesized to reflect mutations in the recep coding and regulatory sequences. THRA function promotes tumor cell development by regulat gene expression of regulatory genes and by influencing metabolic activities (e.g. of key enzyn of alternative metabolic pathways in tumors such as malic enzyme and genes responsible lipogenesis). The observed activities of nuclear receptors not only reflect their transactivat potential, but are also due to posttranscriptional activities in the absence or presence of ligar Co-amplification of THRA /ERBA and ERBB has been shown, but its influence on tur development has been doubted as no overexpression could be demonstrated in breast tumors [de Vijver et al., 1987, (75)]. THRA and RARA are part of nuclear receptor family whose funct can be mediated as monomers, homodimers or heterodimers. RARA regulates differentiation c broad spectrum of cells. Interactions of hormones with ERBB expression has been investigat Ligands of RARA can inhibit the expression of amplified ERBB genes in breast tum [Offterdinger et al., 1998, (76)]. As being part of this invention co-amplification and co-expressior THRA and RARA could be shown. It was also found that multiple genes, which are regulated members of the thyroid hormone receptor - and retinoic acid receptor family, are differentia expressed in tumor samples, corresponding to their genomic alterations (amplification, mutati deletion). These hormone receptor genes and respective target genes are useful to discrimin patient samples with respect to clinical features.

By expression analysis of multiple normal tissues, tumor samples and tumor cell lines a subsequent clustering of the 17q21 region, it was found that the expression profile of Her-2/n positive tumor cells and tumor samples exhibits similarities with the expression pattern of tis from the central nervous system (Fig. 2). This is in line with the observed malformations in central nervous system of Her-2/neu and THRA knock-out mice. Moreover, it was found t NEUROD2, a nuclear factor involved specifically in neurogenesis, is commonly expressed in respective samples. This led to the definition of the 17q21 Locus as being an "ARCHEOI whose primary function in normal organ development is defined to the central nervous syste Surprisingly, the expression of NEUROD2 was affected by therapeutic intervention. Striking also ZNF144, TEM7, PIP5K and PPP1R1B are expressed in neuronal cells, where they disp diverse tissue specific functions.

In addition Her-2/neu is often co-amplified with GRB7, a downstream member of the signal cascade being involved in invasive properties of tumors. Surprisingly, we have found anot member of the Her-2/neu signaling cascade being overexpressed in primary breast tumors TO (= "Transducer of ERBB signaling"). Strong overexpression of TOB1 corellated with wea

overexpression of Her-2/neu, already indicating its involvement in oncogenic signaling activities. Amplification of Her-2/neu has been assigned to enhanced proliferative capacity, due to the identified downstream components of the signaling cascade (e.g. Ras-Raf-MAPK). In this respect it was surprising that some cdc genes, which are cell cycle dependent kinases, are part of the amplicons, which upon altered expression have great impact on cell cycle progression.

The ARCHEONS on 17q21 and 12q13 are very closely related, as they do not only harbor isoforms of specific genes (e.g. CACNB1 vs. CACNB3, ERBB2 vs. ERBB3, RARA vs. RARG, see below), but are even flanked by whole gene clusters, consisting of multiple isoforms of one gene family positioned in tandem, such as the keratin and the HOX gene cluster. In this respect the simultaneous presence of keratins and receptors of the EGFR family, i.e. ERBB2 (Her-2/neu) and ERBB3 (Her-3) is of special interest, as the expression of individual keratins is very tightly controlled by the EGFR signalling pathway.

Keratins are a group of water-insoluble proteins that form 10 nm intermediate filaments in epithelial cells. Approximately 30 different keratin molecules have been identified. They can be divided into acidic and basic-neutral subfamilies according to their relative charges, immunoreactivity, and sequence homologies to types I and II wool keratins, respectively. In vivo, a basic keratin usually is co-expressed and 'paired' with a particular acidic keratin to form a heterodimer. The expression of various keratin pairs is tissue specific, differentiation dependent, and developmentally regulated. The presence of specific keratin pairs is essential for the maintenance of the integrity of epithelium. Alterations of keratin expression have been observed in tumor epithelium, with an abnormal keratin pattern being expressed in tumor cells compared to the adjacent normal tissue. Mutations in human K14/K5 pair and the K10/K1 pair underlie skin diseases such as epidermolysis bullosa simplex and epidermolytic hyperkeratosis. The expression of these and other keratins within the skin is tightly regulated. For example, the expression of K14/K5 pair is restricted to the basal cell layer of the skin displaying no overlap with the K10/K1 pair, which is solely expressed in the suprabasal layer. Gene expression is very tightly controlled by an interplay of multiple signalling cascades such as the EGFR, TGFR, sonic hedgehog and wnt-signaling, involving receptor tyrosine kinases and serin threonin kinases. In addition, posttranslational modifications such as serine/threonine and/or tyrosine phosphorylation events affect keratin function, and can be attributed to receptor tyrosine kinase signalling and MAPK and ERK activity. Posttranlational modifications of keratins not only alters the solubility of keratins, but also affects nuclear and signalling functions (e.g. after association with 14-3-3 protein). In addition, we did observe genomic alteration of the keratin gene clusters perturbing keratin expression pattern.

Moreover, the physical interaction of keratins, which are located in ARCHEONs of differen chromosomes and whose cell type specific expression at distinct differentiation status is regulatec by members of the same ARCHEONs is a superb example of the genetic interaction oi ARCHEON genes. Examples of this tight interaction between the 12q13 and 17q21 ARCHEONs are the expression and physical interaction of keratin 5 (basic keratin Type II located on 12q13) and keratin 14 (acidic keratin Type I located on 17q21) in the basal layer of the skin, which is shut off in the suprabasal layer and compensated by the expression and physical interaction of keratin 1 (basic keratin Type II located on 12q13) and keratin 10 (acidic keratin Type I located on 17q21). Diverse control mechanisms confer this exclusive expression control including chromosomal positioning and growth factor signaling activities. Interestingly, critical keratins are chromosomally postioned in an ordered fashion reflecting their related but exclusive function in different keratin pairs and in specific tissues, resembling the structure and function relationship of the hox gene clusters on the same chromosomes. Moreover, keratins whose mutation result in specific skin disorders (e.g. mutation of K5 and K14 results in hand and foot syndrom) are located at similiar positions within the ARCHEONs on chromosome 17q21 and 12q13. The genes are in close proximity to genes involved in signaling events (e.g. ERBBs and RARs) regulating proliferation, differentiation and apoptotic events also in the skin tissue. For example Her-2/neu is specifically expressed within the basal layer of the skin, where assymmetric cell divisions of adult stem cells or ealry progenitor cells thereof give rise to a non-differentiated daughter cell residing in the basal layer and a differentiating daughter cell which is subsequently moving to the suprabasal compartment. These assymetric cell divisions guarantee the self-renewal and the cellular homeostasis of the skin tissue. This is of importance for the biological functions of the skin such as barrier function towards the environmental stress including infectious agents. Perturbation of the signalling activities within the skin results in diseases similiar to the hereditary disorders reflecting mutations of specific keratin genes. In clinical studies it has been shown, that blocking EGFR signalling by antibody-treatment (e.g. cetuximab) and small molecule inhibitors (e.g. Iressa) targeted to the receptor tyrosin kinases can result in skin diseases (e.g. acne-like rash) of grade I, II or III. It is part of this invention, that these skin diseases not only reflect side effects of the respective treatments, but are an example for systemic changes occuring as a consequence of therapeutic regimen, thereby indicating suscebility of the endogenous signaling network to the therapeutic agents. This observation can have consequences on therapeutic decisions, as the therapeutic regimen are normally stopped or is reduced upon occurence of side effects. However, as the side effects (e.g. the skin dieseases occurring under anti growth factor treatment) are indicative of response to treatment (e.g. tumor shrinkage), the treatment should be endured even though "adverse" drug responses occur and side effects should be treated separately by agents

softening the symptoms. Skin diseases such as rash and hand and foot syndrom are just examples for a given side effect under a given treatment (i.e. anti tumor therapy), that can be used for response correlation.

Similiarly to blocking receptor molecules itself, blocking downstream members of these signaling cascades results mainly in skin diseases (e.g. hand-and-foot syndroms). Surprisingly, we did observe, that treating tumor cells with agents blocking the EGFR/Her-2/neu signaling (e.g. Cetuximab, Iressa, Herceptin, RAF kinase inhibitor, etc.) shifts the expression of specific keratins being part of the ARCHEONS described in this invention. Moreover, the altered expression of keratins in tumor cells of patients is paralleled by a shift of keratin expression in the keratinocytes of the skin of the very same patient. Perturbation of keratin expression and or post-transcriptional modification in the skin tissue seems to resemble the suscebility of the endogenous growth factor signaling pathways to the respective treatment. The resulting skin diseases are therefore at least some extent indicative of the tumor responsiveness to the regimen. This endogenous responsiveness to anti growth factor signaling agents can also be delineated from polymorphisms and genetic alterations (e.g. mutations) being present within the ARCHEON described in this invention. Of particular interest are in this context polymorphisms being present in the keratin genes. However, polymorphisms within keratins, keratin related genes and/or genes functionally connected to the keratin-based cytoskeleton, not necessarily being present within the ARCHEONs described, are also of importance according to their physical interaction with the respective gene products (e.g. ITGB4). It is part of this invention, that the responsiveness of a given tumor to anti growth factor treatment relates to the genetic predisposition of the respective signaling pathway members and target genes, which include keratins and related genes, that are markers for proliferation, differentiation and apoptosis in normal tissues, such as skin tissue. This knowledge can be used to predict the responsiveness of a tumor based on the characterization of surrog tissues, such as skin, blood and any other normal tissue containing the above mentioned genes and/or gene products. For example the responsiveness to Iressa, RAF-kinase inhibitor and antibody based therapies targeting EGFR and Her-2/neu can be delineated from punch biopsies of the skin (preferably by comparison of pre- and/or post-treatment samples) or blood samples by determining the expression pattern or genetic characterization of keratin or keratin-related genes of an individual patient. Moreover, the responsiveness of such surrogate tissues can then be correlated to the tumor phenotype and the responsiveness of a tumor to the respective treatment, thereby predict therapy outcome. The examples of surrogate tissues are given by way of illustration and not by limitation.

It is yet another embodiment of the invention, that adverse drug responses such as heart toxicities can also be deduced from characteristica of the ARCHEONs described. Of particular interest are the ARCEONs at 17q12-24, 12q13 and 3q21-26. It is known that anthracyclin based, anti-cancer regimens result in heart toxicities (such as dilated cardiomyopathies), as can be deduced e.g. by LVEF measurements. Moreover, anthracyclin pretreated patients have significantly increased heart toxicity events upon subsequent Herceptin™ based regimen. Interestingly, the ARCHEONs described in this invention not only harbor the primary targets of these therapies (i.e. topoisomerases and Her-2/neu), but also important structural and functional genes (Telethonin, PNMT, CACNB1, PPARBP, Her-2/neu, Her4) for muscle function including heart muscle function. These genes are involved in central processes of heart muscle function, such as tyrosine phosphorylation, serine/threonine phosphorylation, calcium influx, regulating e.g. central structural proteins such as titin. Moreover, these genes can be colocalized in heart muscles, dispaying their functional interplay in this tissue. In mouse models, the mislocalization of telethonin and the genetic inactivation of Her-2/neu, Her4 and Neuregulin result in a similiar phenotype as can be seen for cancer patients being treated with diverse anti-cancer drugs. The synergistic adverse drug response effect seen for the combinatorial treatment with anthracyclin and Herceptin™. Delineation of polymorphisms and haplotypes of the respective genes, genomic region and/or the ARCHEON structure are indicative of the susceptibility to suffer from heart toxicities upon anti-cancer drug treatment. This is important for therapy decisions and cancer treatment management, as the prior therapies conducted exclude subsequent treatment options. For example, anthracyclin-based pretreatment can exclude subsequent Herceptin™ treatment or lead to reduced dosages, if possible heart toxicities (e.g. dilated cardiomyopathies) cannot be excluded.

According to the observations described above the following examples of genes at 3q21-26 are offered by way of illustration, not by way of limitation.

➔ WNT5A, CACNA1D, THRB, RARB, TOP2B, RAB5B, SMARCC1 (BAF155), RAF, WNT7A

The following examples of genes at 12q13 are offered by way of illustration, not by way of limitation.

➔ CACNB3, Keratins, ERBB3, NR4A1, RAB5/13, RARG, STAT6, WNT10B, (GCN5), (SAS: Sarcoma Amplified Sequence), SMARCC2 (BAF170), SMARCD1 (BAF60A), (GAS41: Glioma Amplified Sequence), (CHOP), Her3, KRTHB, HOX C , IGFBP6, WNT5B

There is cross-talk between the amplified ARCHEONs described above and some other highly amplified genomic regions locate approximately at 1p13, 1q32, 2p16, 2q21, 3p12, 5p13, 6p12, 7p12, 7q21, 8q23, 11q13, 13q12, 19q13, 20q13 and 21q11. The above mentioned chromosomal regions are described by way of illustration not by way of limitation, as the amplified regions often span larger and/or overlapping positions at these chromosomal positions.

Additional alterations of non-transcribed genes, pseudogenes or intergenic regions of said chromosomal locations can be measured for prediction, diagnosis, prognosis, prevention and treatment of malignant neoplasia and breast cancer in particular. Some of the genes or genomic regions have no direct influence on the members of the ARCHEONs or the genes within distinct chromosomal regions but still retain marker gene function due to their chromosomal positioning in the neighborhood of functionally critical genes (e.g. Telethonin neighboring the Her-2/neu gene).

<u>Clinical Relevance of the genes which are part of the 17q21 Archeon for Response to Herceptin treatment</u>

Clinical Samples of patients being treated with Herceptin, docetaxel, paclitaxel, taxotere, carboplatin, cisplatin, oxaliplatin, vinorelbine, tamoxifen, anastrozole, letrozole, tamoxifen, epirubicin, doxorubicin and CMF were obtained. Primary tumor tissues and lymphnode tissues were obtained from neoadjuvant and adjuvant settings. In addition, biopsy material of first and second line therapies was obtained in some cases from metastatic lesions. These samples included formalin-fixed and paraffin-embedded material or fresh tissue from primary tumours and metastatic lesions of the respective patients. Moreover, whole blood, serum and plasma samples were included in the analysis.

Multiparametric, clinical assessment of the response to Herceptin in combination w chemotherapeutics (e.g. docetaxel, taxotere, paclitaxel, vinorelbine, carboplatin, cisplatin), or other therapies described below, was performed, based on clinical information, such as histological parameters (TNM-Stage, AJCC grade), standard molecular markers (IHC staining for estrogen receptor, progesteron receptor, Her-2/neu) and sonographical or radiological assessment (e.g. CT). In addition to combinatorial treatment, samples from single agent therapies were evaluated. Response to treatment was evaluated according to international standards. The ARCHEON genes were analyzed on DNA, RNA or protein level. Normalization of the ARCHEON genes was done by intra- or extrachromosomal reference genes (see EXAMPLE 3 below) or by housekeeping genes of diverse expression level.

We could delineate specific regions of the ARCHEON to be informative for the response tc Herceptin-based therapy. As depicted below, genes that are located towards the centromer telomer of an individual chromosome in relation to a centrally localized gene within ARCHEON (e.g. Her-2/neu in the 17q21 ARCHEON) are in the following named to „centromeric" and „telomeric", respectively. Of particular interest for response to Herceptin-bas treatment are genes being centromeric from the Her-2/neu gene locus on 17q21. The integrity this centromeric ARCHEON region is of importance for the phenotype of Her-2/neu positi tumors. Genetic alteration in the chromosomal region of PIP5K2B, FLJ20291, MLN50, TEM CACNB1, RPL19, MGC15482, PPARBP, CrkRS are critical for clinical outcome of Her-2/ne positive breast tumors. Of particular interest is the centromeric breakpoint region of the 17q2 ARCHEON nearby the genes TEM7, CACNB1, CrkRS and PPARBP. Her-2/neu positive tumoı bearing elevated gene copy numbers of TEM7, CACNB1, CrkRS and PPARBP compared to othe Her-2/neu positive tumors and/or normal tissue controls do have a worsened clinical outcome an a poor response to Herceptin based treatment. The genes within this region are involved in calciur and inositol signalling, which is fundamental with regard to cell survival mechanisms (e.g CACNB1, PPP1R1B and PIP5K2B). Overexpression of CrkRS is of importance for the tumo phenotype, as its kinase activity regultaes the RNA polymerase II holoenzyme complex Especially the phosphorylation of the C-terminal domain and its associated components not onl has influence on the general activity of the enzyme complex, but also affects gene products, whose importance for tumor cell growth has been demonstrated and some of which are part of the ARCHEONs (e.g. the SWI/SNF components SMARCs, e.g. SMARCC2, are critical for RB mediated tumor suppression). Phosphorylation of SMARCs is tightly regulated during cell cycle progression and affects the biological function of the SMARCs (influence on activity, stability and cellular localization). Altered phosphorylation of the RNA polymerase holoenzyme complex by CrkRS therefore most probably affects cell cycle progression. Moreover, the abnormal expression of TEM7, which we found to be elevated in a subclass of Her-2/neu positive tumors, whereas it was originally identified to be a tumor endothelial marker (TEM; see above), points towards an intense interplay between tumor and endothelial cells resulting in a more aggressive behaviour of the respective tumor cells during metastasic processes such as intra- and extravasation. Strikingly, the genes within this region, i.e. ZNF144, TEM7, PIP5K, PPP1R1B and CACNB1, all do have physiological functions within the central nervous system. Therefore, we do assume, that a „neuronal environment" would be favourable for tumor cells overexpressing these genes resulting in growth and survival advantages for these particular tumor cells. In accordance with this, it is observed that Herceptin resistant metastasis frequently occur in the brain. So far it has been discussed, that this observation refers to toxicological problems such as drug-

biovailability with respect to the blood brain barrier. It is part of this invention, that genes which are normally expressed within neuronal cells are integral part of the centromeric gene cluster of the ARCHEON on chromosome 17q21 and are involved in de novo and acquired resistance to Herceptin based treatment. Independent amplification units and/or deletion of singular genes of this centromeric cluster due to chromosomal breakage interferes with the survival and resistance function of this genomic region. Therefore the continuity of amplification units is another important feature with regard to responsivenessor unresponsiveness to therapy. It is noteworthy to mention, that not only the presence of particular genes, but also the presence of regulatory elements within this genomic region contribute to the above mentioned biological phenotype. Therefore also the loss or gain of regulatory elements within the centromeric part of the ARCHEON is of importance for resistance to anti cancer treatment and therefore part of this invention.

In addition to the alteration of centromeric ARCHEON region, the total length of the ARCHEON with regard to the telomeric-region and the relative gene copy numbers of the amplified genes are of importance. Particularly the integrity of the genomic region harboring the TOP2alpha gene with the surrounding genes THRA, NR1D1, MLN51, WIRE, HsCDC6, RARA, CTEN, IGFBP4, EBI1 and SMARCE1 is of interest. Her-2/neu positive tumors, that are deleted in at least some of this genes exhibit a worsened response to herceptin-based chemotherapy. This demonstrates, that this region is not only of prognostic value for anthracyclin-based therapy, but also of prognostic value for chemotherapeutic treatment with taxol-related agents and platin salts. The amplification, deletion or silencing of this telomeric region is accompanied with altered sensitivity to the above mentioned chemotherapeutics. This is a general feature of tumors bearing alterations (with regard to gene expression and/or amplification of the 17q21 ARCHEON) and not only true for breast cancer. In line with this, we have analyzed ovarian tumors bearing alterations in the 17q21 ARCHEON and correlated the clinical outcome, that was assessed similarly as depicted above, with regard to a platin salt based therapy. Strikingly, tumors with defined genetic patterns within this telomeric regions did develop resistance to this chemotherapeutic regimen. Detecting solely the coamplification of Her-2/neu and TOP2alpha was not as informative with regard to response prediction as a detailed characterization and subsequent response correlation with the region of the THRA, NR1D1, MLN51, WIRE, HsCDC6 and RARA genes. It is part of this invention, that the proliferation status of tumors is affected by genes within ARCHEON regions. The 17q21 ARCHEON determines to at least some extent the proliferation rate of tumor cells. Interestingly, Her-2/neu positive tumors bearing elevated levels of a more limited number of genes, excluding several genes in the telomeric region (i.e. TOP2alpha, HsCDC6) exhibit a relatively slow growth rate, which diminishes the effect of chemotherapeutic drugs targeting proliferating cells and is one

of the reasons for the resistance of these tumors to said agents. Instead, these tumors have a higher capacity with regard to invasiveness and do have a diminished apoptotic rate, which to some extent refers to the signaling of Her-2/neu via GRB7 and AKT kinase (also affected by inositols and calcium, see above), respectively. Several genes within the telomeric region of the ARCHEONs affect Her-2/neu signalling, such as RARA, THRA, IGFBP4, and alter the respective characteristics of the cells including proliferation status.

The ARCHEONs being part of this invention, are not only important for clinical response of tumors to antibody-based therapies raised against EGFR- and Her-2/neu signaling (e.g. Herceptin, 2C4 or cetuximab regimen) and to chemotherapeutic agents, but also are of importance for diverse strategies of anti hormonal treatment (e.g. Tamoxifen, Raloxifen, anastrozol, letrozol, faslodex). In particular, elevated levels of the PPARBP gene and protein and the integrity of the telomeric hormone receptor region of the 17q21 ARCHEON, bearing THRA, NR1D1 and RARA, or its related regions on the other ARCHEONs are of importance for these therapeutic regimens. In a retrospective, clinical study evaluating the above mentioned clinical parameters for adjuvant treatment of breast cancer with tamoxifen, we did observe, that the overexpression of PPARBP has impact on the overall survival of patients receiving this therapy. Overexpression of PPARBP enables activity of estrogen and progesteron receptors irrespective of a bound ligand. Therefore, the deregulation of the PPARBP results in the activity of these hormone receptors in the absence of the hormones and even in the presence of anti-hormones and thereby circumvents the anti tumor effect of anti hormonal strategies resulting in resistance of PPARBP overexpressing cells. In addition overexpression of hormone receptors other than estrogen receptor in tumor cells affects activity of estrogen or the respective anti-hormones by competition for dimerization partners, such as RXR, or transcriptional activator or repressor genes, such as CBP or NCOR. With regard to tamoxifen treatment this clearly diminishes the effect of the anti-hormone, as the pool of the transcriptional cofactors is reduced for the classical mode of action of tamoxifen within the nucleus..

The invention further relates to the use of:

A) a polynucleotide comprising at least one of the sequences of SEQ ID NO: 1 to 26 or 53 to 75;

B) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

C) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

D) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c)

E) an antisense molecule targeting specifically one of the polynucleotide sequences specified in (a) to (d);

F) a purified polypeptide encoded by a polynucleotide sequence specified in (a) to (d)

G) a purified polypeptide comprising at least one of the sequences of SEQ ID NO: 27 to 52 or 76 to 98;

H) an antibody capable of binding to one of the polynucleotide specified in (a) to (d) or a polypeptide specified in (f) and (g)

I) a reagent identified by any of the methods of claim 14 to 16 that modulates the amount or activity of a polynucleotide sequence specified in (a) to (d) or a polypeptide specified in (f) and (g)

In the preparation of a composition for the prevention, prediction, diagnosis, prognosis or a medicament for the treatment of malignant neoplasia and breast cancer in particular.

Polynucleotides

A „BREAST CANCER GENE“ polynucleotide can be single- or double-stranded and comprise coding sequence or the complement of a coding sequence for a „BREAST CANCER GENE“ polypeptide. Degenerate nucleotide sequences encoding human „BREAST CANCER GENE“ polypeptides, as well as homologous nucleotide sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, or 98% identical to the nucleotide sequences of SEQ ID NO: 1 to 26or 53 to 75 also are „BREAST CANCER GENE“ polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologues, and variants of „BREAST CANCER GENE“ polynucleotides which encode biologically active „BREAST CANCER GENE“ polypeptides also are „BREAST CANCER GENE“ polynucleotides.

Preparation of Polynucleotides

A naturally occurring „BREAST CANCER GENE" polynucleotide can be isolated free of oth cellular components such as membrane components, proteins, and lipids. Polynucleotides can l made by a cell and isolated using standard nucleic acid purification techniques, or synthesize using an amplification technique, such as the polymerase chain reaction (PCR), or by using a automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the ar Any such technique for obtaining a polynucleotide can be used to obtain isolated „BREAS CANCER GENE" polynucleotides. For example, restriction enzymes and probes can be used t isolate polynucleotide fragments which comprises „BREAST CANCER GENE" nucleotid sequences. Isolated polynucleotides are in preparations which are free or at least 70, 80, or 90% free of other molecules.

„BREAST CANCER GENE" cDNA molecules can be made with standard molecular biolog techniques, using „BREAST CANCER GENE" mRNA as a template. Any RNA isolatio technique which does not select against the isolation of mRNA may be utilized for the purificatio of such RNA samples. See, for example, Sambrook et al., 1989, (77); and Ausubel, F. M. et al. 1989, (78), both of which are incorporated herein by reference in their entirety. Additionally, large numbers of tissue samples may readily be processed using techniques well known to those of skil in the art, such as, for example, the single-step RNA isolation process of Chomczynski, P. (1989, U.S. Pat. No. 4,843,155), which is incorporated herein by reference in its entirety.

„BREAST CANCER GENE" cDNA molecules can thereafter be replicated using molecular biology techniques known in the art and disclosed in manuals such as Sambrook et al., 1989, (77) . An amplification technique, such as PCR, can be used to obtain additional copies of polynucleotides of the invention, using either human genomic DNA or cDNA as a template.

Alternatively, synthetic chemistry techniques can be used to synthesizes „BREAST CANCER GENE" polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a „BREAST CANCER GENE" polypeptide or a biologically active variant thereof.

Identification of differential expression

Transcripts within the collected RNA samples which represent RNA produced by differentially expressed genes may be identified by utilizing a variety of methods which are ell known to those of skill in the art. For example, differential screening [Tedder, T. F. et al., 1988, (79)], subtractive hybridization [Hedrick, S. M. et al., 1984, (80); Lee, S. W. et al., 1984, (81)], and, preferably,

differential display (Liang, P., and Pardee, A. B., 1993, U.S. Pat. No. 5,262,311, which is incorporated herein by reference in its entirety), may be utilized to identify polynucleotide sequences derived from genes that are differentially expressed.

Differential screening involves the duplicate screening of a cDNA library in which one copy of the library is screened with a total cell cDNA probe corresponding to the mRNA population of one cell type while a duplicate copy of the cDNA library is screened with a total cDNA probe corresponding to the mRNA population of a second cell type. For example, one cDNA probe may correspond to a total cell cDNA probe of a cell type derived from a control subject, while the second cDNA probe may correspond to a total cell cDNA probe of the same cell type derived from an experimental subject. Those clones which hybridize to one probe but not to the other potentially represent clones derived from genes differentially expressed in the cell type of interest in control versus experimental subjects.

Subtractive hybridization techniques generally involve the isolation of mRNA taken from two different sources, e.g., control and experimental tissue, the hybridization of the mRNA or single-stranded cDNA reverse-transcribed from the isolated mRNA, and the removal of all hybridized, and therefore double-stranded, sequences. The remaining non-hybridized, single-stranded cDNAs, potentially represent clones derived from genes that are differentially expressed in the two mRNA sources. Such single-stranded cDNAs are then used as the starting material for the construction of a library comprising clones derived from differentially expressed genes.

The differential display technique describes a procedure, utilizing the well known polymerase chain reaction (PCR; the experimental embodiment set forth in Mullis, K. B., 1987, U.S. Pat. No. 4,683,202) which allows for the identification of sequences derived from genes which are differentially expressed. First, isolated RNA is reverse-transcribed into single-stranded cDNA, utilizing standard techniques which are well known to those of skill in the art. Primers for the reverse transcriptase reaction may include, but are not limited to, oligo dT-containing primers, preferably of the reverse primer type of oligonucleotide described below. Next, this technique uses pairs of PCR primers, as described below, which allow for the amplification of clones representing a random subset of the RNA transcripts present within any given cell. Utilizing different pairs of primers allows each of the mRNA transcripts present in a cell to be amplified. Among such amplified transcripts may be identified those which have been produced from differentially expressed genes.

The reverse oligonucleotide primer of the primer pairs may contain an oligo dT stretch of nucleotides, preferably eleven nucleotides long, at its 5' end, which hybridizes to the poly(A) tail

of mRNA or to the complement of a cDNA reverse transcribed from an mRNA poly(A) tai Second, in order to increase the specificity of the reverse primer, the primer may contain one c more, preferably two, additional nucleotides at its 3' end. Because, statistically, only a subset o the mRNA derived sequences present in the sample of interest will hybridize to such primers, th additional nucleotides allow the primers to amplify only a subset of the mRNA derived sequence present in the sample of interest. This is preferred in that it allows more accurate and complet visualization and characterization of each of the bands representing amplified sequences.

The forward primer may contain a nucleotide sequence expected, statistically, to have the abilit to hybridize to cDNA sequences derived from the tissues of interest. The nucleotide sequence ma be an arbitrary one, and the length of the forward oligonucleotide primer may range from about 9 to about 13 nucleotides, with about 10 nucleotides being preferred. Arbitrary primer sequence cause the lengths of the amplified partial cDNAs produced to be variable, thus allowing differen clones to be separated by using standard denaturing sequencing gel electrophoresis. PCR reaction conditions should be chosen which optimize amplified product yield and specificity, and, additionally, produce amplified products of lengths which may be resolved utilizing standard gel electrophoresis techniques. Such reaction conditions are well known to those of skill in the art, and important reaction parameters include, for example, length and nucleotide sequence of oligonucleotide primers as discussed above, and annealing and elongation step temperatures and reaction times. The pattern of clones resulting from the reverse transcription and amplification of the mRNA of two different cell types is displayed via sequencing gel electrophoresis and compared. Differences in the two banding patterns indicate potentially differentially expressed genes.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Randomly-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries can be useful for extension of sequence into 5' nontranscribed regulatory regions.

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate software (e.g. GENOTYPER and Sequence NAVIGATOR,

Perkin Elmer; ABI), and the entire process from loading of samples to computer analysis and electronic data display can be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

Once potentially differentially expressed gene sequences have been identified via bulk techniques such as, for example, those described above, the differential expression of such putatively differentially expressed genes should be corroborated. Corroboration may be accomplished via, for example, such well known techniques as Northern analysis and/or RT-PCR. Upon corroboration, the differentially expressed genes may be further characterized, and may be identified as target and/or marker genes, as discussed, below.

Also, amplified sequences of differentially expressed genes obtained through, for example, differential display may be used to isolate full length clones of the corresponding gene. The full length coding portion of the gene may readily be isolated, without undue experimentation, by molecular biological techniques well known in the art. For example, the isolated differentially expressed amplified fragment may be labeled and used to screen a cDNA library. Alternatively, the labeled fragment may be used to screen a genomic library.

An analysis of the tissue distribution of the mRNA produced by the identified genes may be conducted, utilizing standard techniques well known to those of skill in the art. Such techniques may include, for example, Northern analyses and RT-PCR. Such analyses provide information as to whether the identified genes are expressed in tissues expected to contribute to breast cancer. Such analyses may also provide quantitative information regarding steady state mRNA regulation, yielding data concerning which of the identified genes exhibits a high level of regulation in, preferably, tissues which may be expected to contribute to breast cancer.

Such analyses may also be performed on an isolated cell population of a particular cell type derived from a given tissue. Additionally, standard in situ hybridization techniques may be utilized to provide information regarding which cells within a given tissue express the identified gene. Such analyses may provide information regarding the biological function of an identified gene relative to breast cancer in instances wherein only a subset of the cells within the tissue is thought to be relevant to breast cancer.

Identification of co-amplified genes

Genes involved in genomic alterations (amplifications, insertions, translocations, deletions, etc.) are identified by PCR-based karyotyping in combination with database analysis. Of particular

interest are gene amplifications, which account for gene copy numbers >2 per cell. Gene copy number and gene expression of the respective genes often correlates. Therefore clusters of genes being simultaneously overexpressed due to gene amplifications can be identified by expression analysis via DNA-chip technologies or quantitative RTPCR. For example, the altered expression of genes due to increased or decreased gene copy numbers can be determined by GeneArray™ technologies from Affymetrix or qRT-PCR with the TaqMan or iCycler Systems. Moreover combination of RNA with DNA analytic enables highly parallel and automated characterization of multiple genomic regions of variable length with high resolution in tissue or single cell samples. Furthermore these assays enable the correlation of gene transcription relative to gene copy number of target genes. As there is not necessarily a linear correlation of expression level and gene copy number and as there are synergistic or antagonistic effects in certain gene clusters, the identification on the RNA-level is easier and probably more relevant for the biological outcome of the alterations especially in tumor tissue.

Detection of co-amplified genes in malignant neoplasia

Chromosomal changes are commonly detected by FISH (=Fluorescence-In-Situ-Hybridization) and CGH (=Comparative Genomic Hybridization). For quantification of genomic regions genes or intergenic regions can be used. Such quantification measures the relative abundance of multiple genes with respect to each other (e.g. target gene vs. centromeric region or housekeeping genes). Changes in relative abundance can be detected in paraffin-embedded material even after extraction of RNA or genomic DNA. Measurement of genomic DNA has advantages compared to RNA-analysis due to the stability of DNA, which accounts for the possibility to perform also retrospective studies and offers multiple internal controls (genes not being altered, amplified or deleted) for standardization and exact calculations. Moreover, PCR-analysis of genomic DNA offers the advantage to investigate intergenic, highly variable regions or combinations of SNP's (=Single Nucleotide Polymorphisms), RFLPs, VNTRs and STRs (in general polypmorphic markers). Determination of SNPs or polypmorphic markers within defined genomic regions (e.g. SNP analysis by "Pyrosequencing™") has impact on the phenotype of the genomic alterations. For example it is of advantage to determine combinations of polymorphisms or haplotypes in order to characterize the biological potential of genes being part of amplified alleles. Of particular interest are polypmorphic markers in breakpoint regions, coding regions or regulatory regions of genes or intergenic regions. By determining predictive haplotypes with defined biological or clinical outcome it is possible to establish diagnostic and prognostic assays with non-tumor samples from patients. Depending on whether preferably one allele or both alleles to same extent are amplified (= linear or non-linear amplifications) haplotypes can be determined. Overrepresentation of

specific polypmorphic markers combinations in cells or tissues with gene amplifications facilitates haplotype determination, as e.g. combinations of heterozygous polypmorphic markers in nucleic acids isolated from normal tissues, body fluids or biological samples of one patient become almost homozygous in neoplastic tissue of the very same patient. This "gain of homozygosity" corresponds to the measurement of altered genomic region due to amplification events and is suitable for identification of "gain of function"- alterations in tumors, which result in e.g. oncogenic or growth promoting activities. In contrast, the detection of "losses of heterozygosity" is used for identification of anti-oncogenes, gate keeper genes or checkpoint genes, that suppress oncogenic activities and negatively regulate cellular growth processes. This intrinsic difference clearly opposes the impact of the respective genomic regions for tumor development and emphasizes the significance of "gain of homozygosity" measurements disclosed in this invention. In addition to the analyses on SNPs, a comparative approach of blood leucocyte DNA and tumor DNA based on VNTR detection can reveal the existance of a formerely described ARCHEON. SNP and VNTR sequences and primer sets most suitable for detection of theARCHEON at 17q11-21 are disclosed in Table 4 and Table 6. Detection, quantification and sizing of such polymorphic markers can be achieved by methods known to those with skill in the art. In one embodiment of this invention we disclose the comparative measurement of amount and size of any of the disclosed VNTRs (Table 6) by PCR amplification and capillary electrophoresis. PCR can be carried out by standart protocols favorably in a linear amplification range (low cycle number) and detection by CE should be carried out by suppliers protocols (e.g. Agilent). More favorably the detection of the VNTRs disclosed in Table 6 can be carried out in a multiplex fashion, utilizing a variety of labeled primers (e.g. fluoreszent, radioactive, bioactive) and a suitable CE detection system (e.g. ABI 310). However the detection can also be performed on slab gels consiting of highly concentrated agarose or polyacrylamide with a monochromal DNA stain. Enhancement of resolution can be achieved by appropriate primer design and length variation to give best results in multiplex PCR.

It is also of interest to determine covalent modifications of DNA (e.g. methylation) or the associated chromatin (e.g. acetylation or methylation of associated proteins) within the altered genomic regions, that have impact on transcriptional activity of the genes. In general, by measuring multiple, short sequences (60-300 bp) these techniques enable high-resolution analysis of target regions, which cannot be obtained by conventional methods such as FISH analytic (2-100 kb). Moreover the PCR-based DNA analysis techniques offer advantages with regard to sensitivity, specificity, multiplexing, time consumption and low amount of patient material required. These techniques can be optimized by combination with microdissection or macrodissection to obtain purer starting material for analysis.

Extending Polynucleotides

In one embodiment of such a procedure for the identification and cloning of full length ge sequences, RNA may be isolated, following standard procedures, from an appropriate tissue cellular source. A reverse transcription reaction may then be performed on the RNA using oligonucleotide primer complimentary to the mRNA that corresponds to the amplified fragmer for the priming of first strand synthesis. Because the primer is anti-parallel to the mRN extension will proceed toward the 5' end of the mRNA. The resulting RNA hybrid may then l "tailed" with guanines using a standard terminal transferase reaction, the hybrid may be digeste with RNase H, and second strand synthesis may then be primed with a poly-C primer. Using th two primers, the 5' portion of the gene is amplified using PCR. Sequences obtained may then b isolated and recombined with previously isolated sequences to generate a full-length cDNA of th differentially expressed genes of the invention. For a review of cloning strategies and recombinar DNA techniques, see e.g., Sambrook et al., (77); and Ausubel et al., (78).

Various PCR-based methods can be used to extend the polynucleotide sequences disclosed herei to detect upstream sequences such as promoters and regulatory elements. For example, restrictio site PCR uses universal primers to retrieve unknown sequence adjacent to a known locus [Sarkar 1993, (82)]. Genomic DNA is first amplified in the presence of a primer to a linker sequence and primer specific to the known region. The amplified sequences are then subjected to a second rounc of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region [Triglia et al., 1988 ,(83)]. Primers can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be e.g. 2230 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68-72°C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA [Lagerstrom et al., 1991, (84)]. In this method, multiple restriction enzyme digestions and ligations also can be used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to walk genomic DNA (CLONTECH, Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

The sequences of the identified genes may be used, utilizing standard techniques, to place the genes onto genetic maps, e.g., mouse [Copeland & Jenkins, 1991, (85)] and human genetic maps [Cohen, et al., 1993 ,(86)]. Such mapping information may yield information regarding the genes' importance to human disease by, for example, identifying genes which map near genetic regions to which known genetic breast cancer tendencies map.

Identification of polynucleotide variants and homologues or splice variants

Variants and homologues of the „BREAST CANCER GENE" polynucleotides described above also are „BREAST CANCER GENE" polynucleotides. Typically, homologous „BREAST CANCER GENE" polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known „BREAST CANCER GENE" polynucleotides under stringent conditions, as is known in the art. For example, using the following wash conditions: 2 X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1% SDS, room temperature twice, 30 minutes each; then 2 X SSC, 0.1% SDS, 50 EC once, 30 minutes; then 2 X SSC, room temperature twice, 10 minutes each homologous sequences can be identified which contain at most about 25-30% basepair mismatches. More preferably, homologous polynucleotide strands contain 15-25% basepair mismatches, even more preferably 5-15% basepair mismatches.

Species homologues of the „BREAST CANCER GENE" polynucleotides disclosed herein also can be identified by making suitable probes or primers and screening cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human variants of „BREAST CANCER GENE" polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the T_m of a double-stranded DNA decreases by 1-1.5°C with every 1% decrease in homology [Bonner et al., 1973, (87)]. Variants of human „BREAST CANCER GENE" polynucleotides or „BREAST CANCER GENE" polynucleotides of other species can therefore be identified by hybridizing a putative homologous „BREAST CANCER GENE" polynucleotide with a polynucleotide having a nucleotide sequence of one of the sequences of the SEQ ID NO: 1 to 26 or 53 to 75 or the complement thereof to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to „BREAST CANCER GENE" polynucleotides or their complements following stringent hybridization and/or wash conditions also are „BREAST CANCER GENE" polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example, in Sambrook et al., (77). Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20°C below the calculated T_m of the hybrid under study. The T_m of a hybrid between a „BREAST CANCER GENE" polynucleotide having a nucleotide sequence of one of the sequences of the SEQ ID NO: 1 to 26 or 53 to 75 or the complement thereof and a polynucleotide sequence which is at least about 50, preferably about 75, 90, 96, or 98% identical to one of those nucleotide sequences can be calculated, for example, using the equation below [Bolton and McCarthy, 1962, (88):

$$T_m = 81.5°C - 16.6(\log_{10}[Na^+]) + 0.41(\%G + C) - 0.63(\%formamide) - 600/l),$$

where l = the length of the hybrid in basepairs.

Stringent wash conditions include, for example, 4 X SSC at 65°C, or 50% formamide, 4 X SSC at 28°C, or 0.5 X SSC, 0.1% SDS at 65°C. Highly stringent wash conditions include, for example, 0.2 X SSC at 65°C.

The biological function of the identified genes may be more directly assessed by utilizing relevant in vivo and in vitro systems. In vivo systems may include, but are not limited to, animal systems which naturally exhibit breast cancer predisposition, or ones which have been engineered to exhibit such symptoms, including but not limited to the apoE-deficient malignant neoplasia mouse model [Plump et al., 1992, (89)].

Splice variants derived from the same genomic region, encoded by the same pre mRNA can be identified by hybridization conditions described above for homology search. The specific characteristics of variant proteins encoded by splice variants of the same pre transcript may differ and can also be assayed as disclosed. A „BREAST CANCER GENE" polynucleotide having a nucleotide sequence of one of the sequences of the SEQ ID NO: 1 to 26 or 53 to 75 or the complement thereof may therefor differ in parts of the entire sequence as presented for SEQ ID NO: 60 and the encoded splice variants SEQ ID NO: 61 to 66. These refer to individual proteins SEQ ID NO: 83 to 89. The prediction of splicing events and the identification of the utilized acceptor and donor sites within the pre mRNA can be computed (e.g. Software Package GRAIL or GenomeSCAN) and verified by PCR method by those with skill in the art.

Antisense oligonucleotides

Antisense oligonucleotides are nucleotide sequences which are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form complexes and block either transcription or translation. Preferably, an antisense oligonucleotide is at least 6 nucleotides in length, but can be at least 7, 8, 10, 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a cell as described above to decrease the level of „BREAST CANCER GENE" gene products in the cell.

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, peptide nucleic acids (PNAs; described in U.S. Pat. No. 5,714,331), locked nucleic acids (LNAs; described in WO 99/12826), or a combination of them. Oligonucleotides can be synthesized manually or by an automated synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates, alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamidate, carboxymethyl esters, carbonates, and phosphate triesters[Brown, 1994, (126); Sonveaux, 1994, (127) and Uhlmann et al., 1990, (128)].

Modifications of „BREAST CANCER GENE" expression can be obtained by designing antisense oligonucleotides which will form duplexes to the control, 5', or regulatory regions of the „BREAST CANCER GENE". Oligonucleotides derived from the transcription initiation site, e.g., between positions 10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or chaperons. Therapeutic advances using triplex DNA have been described in the literature [Gee et al., 1994, (129)]. An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Precise complementarity is not required for successful complex formation between an antisense oligonucleotide and the complementary sequence of a „BREAST CANCER GENE" polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to a „BREAST CANCER GENE" polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent „BREAST CANCER GENE" nucleotides, can provide sufficient

targeting specificity for „BREAST CANCER GENE" mRNA. Preferably, each stretch complementary contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in leng Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in length. O skilled in the art can easily use the calculated melting point of an antisense-sense pair to determi the degree of mismatching which will be tolerated between a particular antisense oligonucleoti and a particular „BREAST CANCER GENE" polynucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to hybridize to „BREAST CANCER GENE" polynucleotide. These modifications can be internal or at one both ends of the antisense molecule. For example, internucleoside phosphate linkages can l modified by adding cholesteryl or diamine moieties with varying numbers of carbon residu between the amino groups and terminal ribose. Modified bases and/or sugars, such as arabino instead of ribose, or a 3', 5' substituted oligonucleotide in which the 3' hydroxyl group or the : phosphate group are substituted, also can be employed in a modified antisense oligonucleotid These modified oligonucleotides can be prepared by methods well known in the art[Agrawal e al., 1992, (130); Uhlmann et al., 1987, (131) and Uhlmann et al., (128)].

Ribozymes

Ribozymes are RNA molecules with catalytic activity [Cech, 1987, (132); Cech, 1990, (133) an Couture & Stinchcomb, 1996, (134)]. Ribozymes can be used to inhibit gene function by cleavin an RNA sequence, as is known in the art (e.g., Haseloff et al., U.S. Patent 5,641,673). Th mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecul to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The transcribed sequence of a „BREAST CANCER GENE" can be used to generate ribozymes which will specifically bind to mRNA transcribed from a „BREAST CANCER GENE" genomic locus. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have been developed and described in the art [Haseloff et al., 1988, (135)]. For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target [see, for example, Gerlach et al., EP 0 321201].

Specific ribozyme cleavage sites within a „BREAST CANCER GENE“ RNA target can be identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. Suitability of candidate „BREAST CANCER GENE“ RNA targets also can be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related such that upon hybridizing to the target RNA through the complementary regions, the catalytic region of the ribozyme can cleave the target.

Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease „BREAST CANCER GENE“ expression. Alternatively, if it is desired that the cells stably retain the DNA construct, the construct can be supplied on a plasmid and maintained as a separate element or integrated into the genome of the cells, as is known in the art. A ribozyme-encoding DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

As taught in Haseloff et al., U.S Pat. No. 5,641,673, ribozymes can be engineered so that ribozyme expression will occur in response to factors which induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of regulation, so that destruction of mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

Polypeptides

"BREAST CANCER GENE" polypeptides according to the invention comprise an polypeptic selected from SEQ ID NO: 27 to 52 and 76 to 98 or encoded by any of the polynucleotid sequences of the SEQ ID NO: 1 to 26 and 53 to 75 or derivatives, fragments, analogues an homologues thereof. A "BREAST CANCER GENE" polypeptide of the invention therefore can b a portion, a full-length, or a fusion protein comprising all or a portion of a "BREAST CANCEI GENE" polypeptide.

Protein Purification

„BREAST CANCER GENE" polypeptides can be purified from any cell which expresses th enzyme, including host cells which have been transfected with „BREAST CANCER GENE" expression constructs. Breast tissue is an especially useful source of „BREAST CANCER GENE" polypeptides. A purified „BREAST CANCER GENE" polypeptide is separated from othe compounds which normally associate with the „BREAST CANCER GENE" polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified „BREAST CANCER GENE" polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis.

Obtaining Polypeptides

„BREAST CANCER GENE" polypeptides can be obtained, for example, by purification from human cells, by expression of „BREAST CANCER GENE" polynucleotides, or by direct chemical synthesis.

Biologically Active Variants

„BREAST CANCER GENE" polypeptide variants which are biologically active, i.e., retain an „BREAST CANCER GENE" activity, also are „BREAST CANCER GENE" polypeptides. Preferably, naturally or non-naturally occurring „BREAST CANCER GENE" polypeptide variants have amino acid sequences which are at least about 60, 65, or 70, preferably about 75, 80, 85, 90, 92, 94, 96, or 98% identical to the any of the amino acid sequences of the polypeptides of SEQ ID NO: 27 to 52 or 76 to 98 or the polypeptides encoded by any of the polynucleotides of SEQ ID NO: 1 to 26 or 53 to 75 or a fragment thereof. Percent identity between a putative „BREAST

CANCER GENE" polypeptide variant and of the polypeptides of SEQ ID NO: 27 to 52 or 76 to 98 or the polypeptides encoded by any of the polynucleotides of SEQ ID NO: 1 to 26 or 53 to 75 or a fragment thereof is determined by conventional methods. [See, for example, Altschul *et al.*, 1986, (90 and Henikoff & Henikoff, 1992, (91)]. Briefly, two amino acid sequences are aligned to optimize the alignment scores using a gap opening penalty of 10, a gap extension penalty of 1, and the "BLOSUM62" scoring matrix of Henikoff & Henikoff, (91).

Those skilled in the art appreciate that there are many established algorithms available to align two amino acid sequences. The "FASTA" similarity search algorithm of Pearson & Lipman is a suitable protein alignment method for examining the level of identity shared by an amino acid sequence disclosed herein and the amino acid sequence of a putative variant [Pearson & Lipman, 1988, (92), and Pearson, 1990, (93)]. Briefly, FASTA first characterizes sequence similarity by identifying regions shared by the query sequence (*e.g.*, SEQ ID NO: 1 to 26 or 53 to 75) and a test sequence that have either the highest density of identities (if the ktup variable is 1) or pairs of identities (if ktup=2), without considering conservative amino acid substitutions, insertions, or deletions. The ten regions with the highest density of identities are then rescored by comparing the similarity of all paired amino acids using an amino acid substitution matrix, and the ends of the regions are "trimmed" to include only those residues that contribute to the highest score. If there are several regions with scores greater than the "cutoff" value (calculated by a predetermined formula based upon the length of the sequence the ktup value), then the trimmed initial regions are examined to determine whether the regions can be joined to form an approximate alignment with gaps. Finally, the highest scoring regions of the two amino acid sequences are aligned using a modification of the Needleman-Wunsch-Sellers algorithm [Needleman & Wunsch, 1970, (94), and Sellers, 1974, (95)], which allows for amino acid insertions and deletions. Preferred parameters for FASTA analysis are: ktup=1, gap opening penalty=10, gap extension penalty=1, and substitution matrix=BLOSUM62. These parameters can be introduced into a FASTA program by modifying the scoring matrix file ("SMATRIX"), as explained in Appendix 2 of Pearson, (93).

FASTA can also be used to determine the sequence identity of nucleic acid molecules using a ratio as disclosed above. For nucleotide sequence comparisons, the ktup value can range between one to six, preferably from three to six, most preferably three, with other parameters set as default.

Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity of a „BREAST CANCER GENE" polypeptide can be found using computer programs well known in the art, such as DNASTAR software. Whether an amino acid change results in a biologically active „BREAST CANCER GENE" polypeptide can readily be determined by assaying for „BREAST CANCER GENE" activity, as described for example, in the specific Examples, below. Larger insertions or deletions can also be caused by alternative splicing. Protein domains can be inserted or deleted without altering the main activity of the protein.

Fusion Proteins

Fusion proteins are useful for generating antibodies against „BREAST CANCER GENE" polypeptide amino acid sequences and for use in various assay systems. For example, fusion proteins can be used to identify proteins which interact with portions of a „BREAST CANCER GENE" polypeptide. Protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A „BREAST CANCER GENE" polypeptide fusion protein comprises two polypeptide segments fused together by means of a peptide bond. The first polypeptide segment comprises at least 25, 50, 75, 100, 150, 200, 300, 400, 500, 600, 700 or 750 contiguous amino acids of an amino acid sequence encoded by any polynucleotide sequences of the SEQ ID NO: 1 to 26 or 53 to 75 or of a biologically active variant, such as those described above. The first polypeptide segment also can comprise full-length „BREAST CANCER GENE".

The second polypeptide segment can be a full-length protein or a protein fragment. Proteins commonly used in fusion protein construction include β-galactosidase, β-glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horseradish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A fusion protein also can be engineered to contain a cleavage site located between the „BREAST CANCER GENE" polypeptide-encoding

sequence and the heterologous protein sequence, so that the „BREAST CANCER GENE“ polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two polypeptide segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises coding sequences selected from any of the polynucleotide sequences of the SEQ ID NO: 1 to 26 and 53 to 75 in proper reading frame with nucleotides encoding the second polypeptide segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

Identification of Species Homologues

Species homologues of human a „BREAST CANCER GENE“ polypeptide can be obtained using „BREAST CANCER GENE“ polypeptide polynucleotides (described below) to make suitable probes or primers for screening cDNA expression libraries from other species, such as mice, monkeys, or yeast, identifying cDNAs which encode homologues of a „BREAST CANCER GENE“ polypeptide, and expressing the cDNAs as is known in the art.

Expression of Polynucleotides

To express a „BREAST CANCER GENE“ polynucleotide, the polynucleotide can be inserted into an expression vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding „BREAST CANCER GENE“ polypeptides and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described, for example, in Sambrook et al., (77) and in Ausubel et al., (78).

A variety of expression vector/host systems can be utilized to contain and express sequences encoding a „BREAST CANCER GENE“ polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors, insect cell systems

infected with virus expression vectors (e.g., baculovirus), plant cell systems transformed wi virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) with bacterial expression vectors (e.g., Ti or pBR322 plasmids), or animal cell systems.

The control elements or regulatory sequences are those regions of the vector enhancers, promoter 5' and 3' untranslated regions which interact with host cellular proteins to carry out transcriptic and translation. Such elements can vary in their strength and specificity. Depending on the vecto system and host utilized, any number of suitable transcription and translation elements, includin constitutive and inducible promoters, can be used. For example, when cloning in bacterial system: inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. Th baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived fror the genomes of plant cells (e.g., heat shock, RUBISCO, and storage protein genes) or from plan viruses (e.g., viral promoters or leader sequences) can be cloned into the vector. In mammalia cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it i: necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a „BREAST CANCER GENE" polypeptide, vectors based on SV40 or EBV can be used with ar appropriate selectable marker.

Bacterial and Yeast Expression Systems

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the „BREAST CANCER GENE“ polypeptide. For example, when a large quantity of the „BREAST CANCER GENE“ polypeptide is needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene). In a BLUESCRIPT vector, a sequence encoding the „BREAST CANCER GENE“ polypeptide can be ligated into the vector in frame with sequences for the amino terminal Met and the subsequent 7 residues of ß-galactosidase so that a hybrid protein is produced. pIN vectors [Van Heeke & Schuster, (17)] or pGEX vectors (Promega, Madison, Wis.) also can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast Saccharomyces cerevisiae, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used. For reviews, see Ausubel et al., (4) and Grant et al., (18).

Plant and Insect Expression Systems

If plant expression vectors are used, the expression of sequences encoding „BREAST CANCER GENE“ polypeptides can be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV can be used alone or in combination with the omega leader sequence from TMV [Takamatsu, 1987, (96)]. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters can be used [Coruzzi et al., 1984, (97); Broglie et al., 1984, (98); Winter et al., 1991, (99)]. These constructs can be introduced into plant cells by direct DNA transformation or by pathogen-mediated transfection. Such techniques are described in a number of generally available reviews.

An insect system also can be used to express a „BREAST CANCER GENE“ polypeptide. For example, in one such system Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in Spodoptera frugiperda cells or in Trichoplusia larvae. Sequences encoding „BREAST CANCER GENE“ polypeptides can be cloned into a nonessential

region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of „BREAST CANCER GENE" polypeptides will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can then be used to infect S. frugiperda cells or Trichoplusia larvae in which „BREAST CANCER GENE" polypeptides can be expressed [Engelhard et al., 1994, (100)].

Mammalian Expression Systems

A number of viral-based expression systems can be used to express „BREAST CANCER GENE" polypeptides in mammalian host cells. For example, if an adenovirus is used as an expression vector, sequences encoding „BREAST CANCER GENE" polypeptides can be ligated into an adenovirus transcription/translation complex comprising the late promoter and tripartite leader sequence. Insertion in a nonessential E1 or E3 region of the viral genome can be used to obtain a viable virus which is capable of expressing a „BREAST CANCER GENE" polypeptide in infected host cells [Logan & Shenk, 1984, (101)]. If desired, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, can be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are constructed and delivered to cells via conventional delivery methods (e.g., liposomes, polycationic amino polymers, or vesicles).

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding „BREAST CANCER GENE" polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a „BREAST CANCER GENE" polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used [Scharf et al., 1994, (102)].

Host Cells

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed „BREAST CANCER GENE" polypeptide in the desired fashion. Such

modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Posttranslational processing which cleaves a "prepro" form of the polypeptide also can be used to facilitate correct insertion, folding and/or function. Different host cells which have specific cellular machinery and characteristic mechanisms for Post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express „BREAST CANCER GENE“ polypeptides can be transformed using expression vectors which can contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vect[illegible] Following the introduction of the vector, cells can be allowed to grow for 12 days in an enriched medium before they are switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced „BREAST CANCER GENE“ sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type [Freshney et al., 1986, (103).

Any number of selection systems can be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler et al., 1977, (104)] and adenine phosphoribosyltransferase [Lowy et al., 1980, (105)] genes which can be employed in tk^- or $aprt^-$ cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, dhfr confers resistance to methotrexate [Wigler et al., 19[illegible] (106)], npt confers resistance to the aminoglycosides, neomycin and G418 [Colbere-Garapin et al., 1981, (107)], and als and pat confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. Additional selectable genes have been described. For example, trpB allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine [Hartman & Mulligan, 1988 ,(108)]. Visible markers such as anthocyanins, ß-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system [Rhodes et al., 1995, (109)].

Detecting Expression and gene product

Although the presence of marker gene expression suggests that the „BREAST CANCER GENE polynucleotide is also present, its presence and expression may need to be confirmed. Fc example, if a sequence encoding a „BREAST CANCER GENE" polypeptide is inserted within marker gene sequence, transformed cells containing sequences which encode a „BREAS CANCER GENE" polypeptide can be identified by the absence of marker gene functior Alternatively, a marker gene can be placed in tandem with a sequence encoding a „BREAS CANCER GENE" polypeptide under the control of a single promoter. Expression of the marke gene in response to induction or selection usually indicates expression of the „BREAST CANCEI GENE" polynucleotide.

Alternatively, host cells which contain a „BREAST CANCER GENE" polynucleotide and whicl express a „BREAST CANCER GENE" polypeptide can be identified by a variety of procedure known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA o DNA-RNA hybridization and protein bioassay or immunoassay techniques which include membrane, solution, or chip-based technologies for the detection and/or quantification o polynucleotide or protein. For example, the presence of a polynucleotide sequence encoding a „BREAST CANCER GENE" polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding a „BREAST CANCER GENE" polypeptide. Nucleic acid amplification-based assays involve the use of oligonucleotides selected from sequences encoding a „BREAST CANCER GENE" polypeptide to detect transformants which contain a „BREAST CANCER GENE" polynucleotide.

A variety of protocols for detecting and measuring the expression of a „BREAST CANCER GENE" polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a „BREAST CANCER GENE" polypeptide can be used, or a competitive binding assay can be employed. These and other assays are described in Hampton et al., (110) and Maddox et al., 111).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding „BREAST CANCER GENE" polypeptides include oligo labeling, nick translation, end-labeling, or PCR amplification

using a labeled nucleotide. Alternatively, sequences encoding a „BREAST CANCER GENE" polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes in vitro by addition of labeled nucleotides and an appropriate RNA polymerase such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, and fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Expression and Purification of Polypeptides

Host cells transformed with nucleotide sequences encoding a „BREAST CANCER GENE" polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or stored intracellular depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode „BREAST CANCER GENE" polypeptides can be designed to contain signal sequences which direct secretion of soluble „BREAST CANCER GENE" polypeptides through a prokaryotic or eukaryotic cell membrane or which direct the membrane insertion of membrane-bound „BREAST CANCER GENE" polypeptide.

As discussed above, other constructions can be used to join a sequence encoding a „BREAST CANCER GENE" polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). Inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the „BREAST CANCER GENE" polypeptide also can be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a „BREAST CANCER GENE" polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilized metal ion affinity chromatography [Porath et al., 1992, (112)], while the enterokinase cleavage site provides a means for purifying the „BREAST CANCER GENE" polypeptide from the fusion protein. Vectors which contain fusion proteins are disclosed in Kroll et al., (113).

Chemical Synthesis

Sequences encoding a „BREAST CANCER GENE“ polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers et al., (114) and Horn et al. (115). Alternatively, a „BREAST CANCER GENE“ polypeptide itself can be produced using chemical methods to synthesize its amino acid sequence, such as by direct peptide synthesis using solid-phase techniques [Merrifield, 1963, (116) and Roberge et al., 1995, (117)]. Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Optionally, fragments of „BREAST CANCER GENE“ polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography [Creighton, 1983, (118)]. The composition of a synthetic „BREAST CANCER GENE“ polypeptide can be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure; see Creighton, (118). Additionally, any portion of the amino acid sequence of the „BREAST CANCER GENE“ polypeptide can be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a fusion protein.

Production of Altered Polypeptides

As will be understood by those of skill in the art, it may be advantageous to produce „BREAST CANCER GENE“ polypeptide-encoding nucleotide sequences possessing non-natural occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter „BREAST CANCER GENE“ polypeptide-encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the polypeptide or mRNA product. DNA shuffling by random fragmentation and PCR re-assembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

Predictive, Diagnostic and Prognostic Assays

The present invention provides method for determining whether a subject is at risk for developing malignant neoplasia and breast cancer in particular by detecting one of the disclosed polynucleotide markers comprising any of the polynucleotides sequences of the SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19 or 21 to 26 or 53 to 75 and/or the polypeptide markers encoded thereby or polypeptide markers comprising any of the polypeptide sequences of the SEQ ID NO: 28 to 32, 34, 35, 37 to 42, 44, 45 or 47 to 52 or 76 to 98 or at least 2 of the disclosed polynucleotides selected from SEQ ID NO: 1 to 26 and 53 to 75 or the at least 2 of the disclosed polypeptides selected from SEQ ID NO: 28 to 32 and 76 to 98 for malignant neoplasia and breast cancer in particular.

In clinical applications, biological samples can be screened for the presence and/or absence of the biomarkers identified herein. Such samples are for example needle biopsy cores, surgical resection samples, or body fluids like serum, thin needle nipple aspirates and urine. For example, these methods include obtaining a biopsy, which is optionally fractionated by cryostat sectioning to enrich diseases cells to about 80% of the total cell population. In certain embodiments, polynucleotides extracted from these samples may be amplified using techniques well known in the art. The expression levels of selected markers detected would be compared with statistically valid groups of diseased and healthy samples.

In one embodiment the diagnostic method comprises determining whether a subject has an abnormal mRNA and/or protein level of the disclosed markers, such as by Northern blot analysis, reverse transcription-polymerase chain reaction (RT-PCR), in situ hybridization, immunoprecipitation, Western blot hybridization, or immunohistochemistry. According to the method, cells are obtained from a subject and the levels of the disclosed biomarkers, protein or mRNA level, is determined and compared to the level of these markers in a healthy subject. An abnormal level of the biomarker polypeptide or mRNA levels is likely to be indicative of malignant neoplasia such as breast cancer.

In another embodiment the diagnostic method comprises determining whether a subject has an abnormal DNA content of said genes or said genomic loci, such as by Southern blot analysis, dot blot analysis, fluorescence or colorimetric In Situ hybridization, comparative genomic hybridization, genotpying by VNTR, STS-PCR or quantitative PCR. In general these assays comprise the usage of probes from representative genomic regions. The probes contain at least parts of said genomic regions or sequences complementary or analogous to said regions. In particular intra- or intergenic regions of said genes or genomic regions. The probes can consist of

nucleotide sequences or sequences of analogous functions (e.g. PNAs, Morpholino oligomers) being able to bind to target regions by hybridization. In general genomic regions being altered in said patient samples are compared with unaffected control samples (normal tissue from the same or different patients, surrounding unaffected tissue, peripheral blood) or with genomic regions of the same sample that don't have said alterations and can therefore serve as internal controls. In a preferred embodiment regions located on the same chromosome are used. Alternatively, gonosomal regions and /or regions with defined varying amount in the sample are used. In one favored embodiment the DNA content, structure, composition or modification is compared that lie within distinct genomic regions. Especially favored are methods that detect the DNA content of said samples, where the amount of target regions are altered by amplification and or deletions. In another embodiment the target regions are analyzed for the presence of polymorphisms (e.g. Single Nucleotide Polymorphisms or mutations) that affect or predispose the cells in said samples with regard to clinical aspects, being of diagnostic, prognostic or therapeutic value. Preferably, the identification of sequence variations is used to define haplotypes that result in characteristic behavior of said samples with said clinical aspects.

The following examples of genes in 17q12-21.2 are offered by way of illustration, not by way of limitation.

One embodiment of the invention is a method for the prediction, diagnosis or prognosis of malignant neoplasia by the detection of at least10, at least 5, or at least 4, or at least 3 and more preferably at least 2 markers whereby the markers are genes and fragments thereof and/or genomic nucleic acid sequences that are located on one chromosomal region which is altered in malignant neoplasia.

One further embodiment of the invention is method for the prediction, diagnosis or prognosis of malignant neoplasia by the detection of at least 10, at least 5, or at least 4, or at least 3 and more preferably at least 2 markers whereby the markers (a) are genes and fragments thereof and/or genomic nucleic acid sequences that are located on one or more chromosomal region(s) which is/are altered in malignant neoplasia and (b) functionally interact as (i) receptor and ligand or (ii) members of the same signal transduction pathway or (iii)members of synergistic signal transduction pathways or (iv) members of antagonistic signal transduction pathways or (v) transcription factor and transcription factor binding site.

In one embodiment, the method for the prediction, diagnosis or prognosis of malignant neoplasia and breast cancer in particular is done by the detection of:

(a) polynucleotide selected from the polynucleotides of the SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26 or 53 to 75;

(b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3;

(c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3;

(d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

in a biological sample comprising the following steps: hybridizing any polynucleotide or analogous oligomer specified in (a) to (do) to a polynucleotide material of a biological sample, thereby forming a hybridization complex; and detecting said hybridization complex.

In another embodiment the method for the prediction, diagnosis or prognosis of malignant neoplasia is done as just described but, wherein before hybridization, the polynucleotide material of the biological sample is amplified.

In another embodiment the method for the diagnosis or prognosis of malignant neoplasia and breast cancer in particular is done by the detection of:

(a) a polynucleotide selected from the polynucleotides of the SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26 or 53 to 75;

(b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3;

(c) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3;

(d) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

(e) a polypeptide encoded by a polynucleotide sequence specified in (a) to (d)

(f) a polypeptide comprising any polypeptide of SEQ ID NO: 28 to 32, 34, 35, 37 to 42, 4 45, 47 to 52 or 76 to 98;

comprising the steps of contacting a biological sample with a reagent which specifically interac with the polynucleotide specified in (a) to (d) or the polypeptide specified in (e).

DNA array technology

In one embodiment, the present Invention also provides a method wherein polynucleotide prob are immobilized an a DNA chip in an organized array. Oligonucleotides can be bound to a soli Support by a variety of processes, including lithography. For example a chip can hold up t 4100,00 oligonucleotides (GeneChip, Affymetrix). The present invention provides significa advantages over the available tests for malignant neoplasia, such as breast cancer, because increases the reliability of the test by providing an array of polynucleotide markers an a singl chip.

The method includes obtaining a biopsy of an affected person, which is optionally fractionated b cryostat sectioning to enrich diseased cells to about 80% of the total cell population and the use o body fluids such as serum or urine, serum or cell containing liquids (e.g. derived from fine needl aspirates). The DNA or RNA is then extracted, amplified, and analyzed with a DNA chip to determine the presence of absence of the marker polynucleotide sequences. In one embodiment the polynucleotide probes are spotted onto a substrate in a two-dimensional matrix or array. samples of polynucleotides can be labeled and then hybridized to the probes. Double-stranded polynucleotides, comprising the labeled sample polynucleotides bound to probe polynucleotides, can be detected once the unbound portion of the sample is washed away.

The probe polynucleotides can be spotted an substrates including glass, nitrocellulose, etc. The probes can be bound to the Substrate by either covalent bonds or by non-specific interactions, such as hydrophobic interactions. The sample polynucleotides can be labeled using radioactive labels, fluorophores, chromophores, etc. Techniques for constructing arrays and methods of using these arrays are described in EP 0 799 897; WO 97/29212; WO 97/27317; EP 0 785 280; WO 97/02357; U.S. Pat. No. 5,593,839; U.S. Pat. No. 5,578,832; EP 0 728 520; U.S. Pat. No. 5,599,695; EP 0 721 016; U.S. Pat. No. 5,556,752; WO 95/22058; and U.S. Pat. No. 5,631,734. Further, arrays can be used to examine differential expression of genes and can be used to determine gene function. For example, arrays of the instant polynucleotide sequences can be used to determine if any of the polynucleotide sequences are differentially expressed between normal cells and diseased cells, for

example. High expression of a particular message in a diseased sample, which is not observed in a corresponding normal sample, can indicate a breast cancer specific protein.

Accordingly, in one aspect, the invention provides probes and primers that are specific to the unique polynucleotide markers disclosed herein.

In one embodiment, the method comprises using a polynucleotide probe to determine the presence of malignant or breast cancer cells in particular in a tissue from a patient. Specifically, the method comprises:

1) providing a polynucleotide probe comprising a nucleotide sequence at least 12 nucleotides in length, preferably at least 15 nucleotides, more preferably, 25 nucleotides, and most preferably at least 40 nucleotides, and up to all or nearly all of the coding sequence which is complementary to a portion of the coding sequence of a polynucleotide selected from the polynucleotides of SEQ ID NO: 1 to 26 and 53 to 75 or a sequence complementary thereto and is

2) differentially expressed in malignant neoplasia, such as breast cancer;

3) obtaining a tissue sample from a patient with malignant neoplasia;

4) providing a second tissue sample from a patient with no malignant neoplasia;

5) contacting the polynucleotide probe under stringent conditions with RNA of each of said first and second tissue samples (e.g., in a Northern blot or in situ hybridization assay); and

6) comparing (a) the amount of hybridization of the probe with RNA of the first tissue sample, with (b) the amount of hybridization of the probe with RNA of the second tissue sample;

wherein a statistically significant difference in the amount of hybridization with the RNA of the first tissue sample as compared to the amount of hybridization with the RNA of the second tissue sample is indicative of malignant neoplasia and breast cancer in particular in the first tissue sample.

Data analysis methods

Comparison of the expression levels of one or more "BREAST CANCER GENES" with reference expression levels, e.g., expression levels in diseased cells of breast cancer or in normal counterpart cells, is preferably conducted using computer systems. In one embodiment, expression levels are

obtained in two cells and these two sets of expression levels are introduced into a computer systei for comparison. In a preferred embodiment, one set of expression levels is entered into a comput‹ system for comparison with values that are already present in the computer system, or i computer-readable form that is then entered into the computer system.

In one embodiment, the invention provides a computer readable form of the gene expressio profile data of the invention, or of values corresponding to the level of expression of at least on "BREAST CANCER GENE" in a diseased cell. The values can be mRNA expression level obtained from experiments, e.g., microarray analysis. The values can also be mRNA level normalised relative to a reference gene whose expression is constant in numerous cells unde numerous conditions, e.g., GAPDH. In other embodiments, the values in the computer are ratio: of, or differences between, normalized or non-normalized mRNA levels in different samples.

The gene expression profile data can be in the form of a table, such as an Excel table. The data car be alone, or it can be part of a larger database, e.g., comprising other expression profiles. Foı example, the expression profile data of the invention can be part of a public database. The computer readable form can be in a computer. In another embodiment, the invention provides a computer displaying the gene expression profile data.

In one embodiment, the invention provides a method for determining the similarity between the level of expression of one or more "BREAST CANCER GENES" in a first cell, e.g., a cell of a subject, and that in a second cell, comprising obtaining the level of expression of one or more "BREAST CANCER GENES" in a first cell and entering these values into a computer comprising a database including records comprising values corresponding to levels of expression of one or more "BREAST CANCER GENES" in a second cell, and processor instructions, e.g., a user interface, capable of receiving a selection of one or more values for comparison purposes with data that is stored in the computer. The computer may further comprise a means for converting the comparison data into a diagram or chart or other type of output.

In another embodiment, values representing expression levels of "BREAST CANCER GENES" are entered into a computer system, comprising one or more databases with reference expression levels obtained from more than one cell. For example, the computer comprises expression data of diseased and normal cells. Instructions are provided to the computer, and the computer is capable of comparing the data entered with the data in the computer to determine whether the data entered is more similar to that of a normal cell or of a diseased cell.

In another embodiment, the computer comprises values of expression levels in cells of subjects at different stages of breast cancer, and the computer is capable of comparing expression data entered into the computer with the data stored, and produce results indicating to which of the expression profiles in the computer, the one entered is most similar, such as to determine the stage of breast cancer in the subject.

In yet another embodiment, the reference expression profiles in the computer are expression profiles from cells of breast cancer of one or more subjects, which cells are treated *in vivo* or *in vitro* with a drug used for therapy of breast cancer.. Upon entering of expression data of a cell of a subject treated *in vitro* or *in vivo* with the drug, the computer is instructed to compare the data entered to the data in the computer, and to provide results indicating whether the expression data input into the computer are more similar to those of a cell of a subject that is responsive to the drug or more similar to those of a cell of a subject that is not responsive to the drug. Thus, the results indicate whether the subject is likely to respond to the treatment with the drug or unlikely to respond to it.

In one embodiment, the invention provides a system that comprises a means for receiving gene expression data for one or a plurality of genes; a means for comparing the gene expression data from each of said one or plurality of genes to a common reference frame; and a means for presenting the results of the comparison. This system may further comprise a means for clustering the data.

In another embodiment, the invention provides a computer program for analyzing gene expression data comprising (i) a computer code that receives as input gene expression data for a plurality of genes and (ii) a computer code that compares said gene expression data from each of said plurality of genes to a common reference frame.

The invention also provides a machine-readable or computer-readable medium including program instructions for performing the following steps: (i) comparing a plurality of values corresponding to expression levels of one or more genes characteristic of breast cancer in a query cell with a database including records comprising reference expression or expression profile data of one or more reference cells and an annotation of the type of cell; and (ii) indicating to which cell the query cell is most similar based on similarities of expression profiles. The reference cells can be cells from subjects at different stages of breast cancer. The reference cells can also be cells from subjects responding or not responding to a particular drug treatment and optionally incubated *in vitro* or *in vivo* with the drug.

The reference cells may also be cells from subjects responding or not responding to seve different treatments, and the computer system indicates a preferred treatment for the subje Accordingly, the invention provides a method for selecting a therapy for a patient having bre: cancer, the method comprising: (i) providing the level of expression of one or more gen characteristic of breast cancer in a diseased cell of the patient; (ii) providing a plurality reference profiles, each associated with a therapy, wherein the subject expression profile and ea reference profile has a plurality of values, each value representing the level of expression of a ge characteristic of breast cancer; and (iii) selecting the reference profile most similar to the subje expression profile, to thereby select a therapy for said patient. In a preferred embodiment step (ii is performed by a computer. The most similar reference profile may be selected by weighing comparison value of the plurality using a weight value associated with the correspondin expression data.

The relative abundance of an mRNA in two biological samples can be scored as a perturbation an its magnitude determined (i.e., the abundance is different in the two sources of mRNA tested), o as not perturbed (i.e., the relative abundance is the same). In various embodiments, a differenc between the two sources of RNA of at least a factor of about 25% (RNA from one source is 25% more abundant in one source than the other source), more usually about 50%, even more often b a factor of about 2 (twice as abundant), 3 (three times as abundant) or 5 (five times as abundant) i scored as a perturbation. Perturbations can be used by a computer for calculating and expressior comparisons.

Preferably, in addition to identifying a perturbation as positive or negative, it is advantageous to determine the magnitude of the perturbation. This can be carried out, as noted above, by calculating the ratio of the emission of the two fluorophores used for differential labeling, or by analogous methods that will be readily apparent to those of skill in the art.

The computer readable medium may further comprise a pointer to a descriptor of a stage of breast cancer or to a treatment for breast cancer.

In operation, the means for receiving gene expression data, the means for comparing the gene expression data, the means for presenting, the means for normalizing, and the means for clustering within the context of the systems of the present invention can involve a programmed computer with the respective functionalities described herein, implemented in hardware or hardware and software; a logic circuit or other component of a programmed computer that performs the operations specifically identified herein, dictated by a computer program; or a computer memory

encoded with executable instructions representing a computer program that can cause a computer to function in the particular fashion described herein.

Those skilled in the art will understand that the systems and methods of the present invention may be applied to a variety of systems, including IBM-compatible personal computers running MS-DOS or Microsoft Windows.

The computer may have internal components linked to external components. The internal components may include a processor element interconnected with a main memory. The computer system can be an Intel Pentium®-based processor of 200 MHz or greater clock rate and with 32 MB or more of main memory. The external component may comprise a mass storage, which can be one or more hard disks (which are typically packaged together with the processor and memory). Such hard disks are typically of 1 GB or greater storage capacity. Other external components include a user interface device, which can be a monitor, together with an inputing device, which can be a "mouse", or other graphic input devices, and/or a keyboard. A printing device can also be attached to the computer.

Typically, the computer system is also linked to a network link, which can be part of an Ethernet link to other local computer systems, remote computer systems, or wide area communication networks, such as the Internet. This network link allows the computer system to share data and processing tasks with other computer systems.

Loaded into memory during operation of this system are several software components, which are both standard in the art and special to the instant invention. These software components collectively cause the computer system to function according to the methods of this invention. These software components are typically stored on a mass storage. A software component represents the operating system, which is responsible for managing the computer system and its network interconnections. This operating system can be, for example, of the Microsoft Windows' family, such as Windows 95, Windows 98, or Windows NT. A software component represents common languages and functions conveniently present on this system to assist programs implementing the methods specific to this invention. Many high or low level computer languages can be used to program the analytic methods of this invention. Instructions can be interpreted during run-time or compiled. Preferred languages include C/C++, and JAVA®. Most preferably, the methods of this invention are programmed in mathematical software packages which allow symbolic entry of equations and high-level specification of processing, including algorithms to be used, thereby freeing a user of the need to procedurally program individual equations or algorithms. Such packages include Matlab from Mathworks (Natick, Mass.), Mathematica from

Wolfram Research (Champaign, Ill.), or S-Plus from Math Soft (Cambridge, Mass.). Accordingly a software component represents the analytic methods of this invention as programmed in procedural language or symbolic package. In a preferred embodiment, the computer system als contains a database comprising values representing levels of expression of one or more gene characteristic of breast cancer. The database may contain one or more expression profiles of gene characteristic of breast cancer in different cells.

In an exemplary implementation, to practice the methods of the present invention, a user firs loads expression profile data into the computer system. These data can be directly entered by th user from a monitor and keyboard, or from other computer systems linked by a networl connection, or on removable storage media such as a CD-ROM or floppy disk or through th network. Next the user causes execution of expression profile analysis software which perform the steps of comparing and, e.g., clustering co-varying genes into groups of genes.

In another exemplary implementation, expression profiles are compared using a method describec in U.S. Patent No. 6,203,987. A user first loads expression profile data into the computer system. Geneset profile definitions are loaded into the memory from the storage media or from a remote computer, preferably from a dynamic geneset database system, through the network. Next the user causes execution of projection software which performs the steps of converting expression profile to projected expression profiles. The projected expression profiles are then displayed.

In yet another exemplary implementation, a user first leads a projected profile into the memory. The user then causes the loading of a reference profile into the memory. Next, the user causes the execution of comparison software which performs the steps of objectively comparing the profiles.

Detection of variant polynucleotide sequence

In yet another embodiment, the invention provides methods for determining whether a subject is at risk for developing a disease, such as a predisposition to develop malignant neoplasia, for example breast cancer, associated with an aberrant activity of any one of the polypeptides encoded by any of the polynucleotides of the SEQ ID NO: 1 to 26 or 53 to 75, wherein the aberrant activity of the polypeptide is characterized by detecting the presence or absence of a genetic lesion characterized by at least one of these:

(i) an alteration affecting the integrity of a gene encoding a marker polypeptides, or

(ii) the misexpression of the encoding polynucleotide.

To illustrate, such genetic lesions can be detected by ascertaining the existence of at least one of these:

I. a deletion of one or more nucleotides from the polynucleotide sequence

II. an addition of one or more nucleotides to the polynucleotide sequence

III. a substitution of one or more nucleotides of the polynucleotide sequence

IV. a gross chromosomal rearrangement of the polynucleotide sequence

V. a gross alteration in the level of a messenger RNA transcript of the polynucleotide sequence

VI. aberrant modification of the polynucleotide sequence, such as of the methylation pattern of the genomic DNA

VII. the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene

VIII. a non-wild type level of the marker polypeptide

IX. allelic loss of the gene

X. allelic gain of the gene

XI. inappropriate post-translational modification of the marker polypeptide

The present Invention provides assay techniques for detecting mutations in the encoding polynucleotide sequence. These methods include, but are not limited to, methods involving sequence analysis, Southern blot hybridization, restriction enzyme site mapping, and methods involving detection of absence of nucleotide pairing . between the polynucleotide to be analyzed and a probe.

Specific diseases or disorders, e.g., genetic diseases or disorders, are associated with specific allelic variants of polymorphic regions of certain genes, which do not necessarily encode a mutated protein. Thus, the presence of a specific allelic variant of a polymorphic region of a gene in a subject can render the subject susceptible to developing a specific disease or disorder. Polymorphic regions in genes, can be identified, by determining the nucleotide sequence of genes in populations of individuals. If a polymorphic region is identified, then the link with a specific disease can be determined by studying specific populations of individuals, e.g. individuals which developed a specific disease, such as breast cancer. A polymorphic region can be located in any

region of a gene, e.g., exons, in coding or non coding regions of exons, introns, and promote region.

In an exemplary embodiment, there is provided a polynucleotide composition comprising a polynucleotide probe including a region of nucleotide sequence which is capable of hybridising to a sense or antisense sequence of a gene or naturally occurring mutants thereof, or 5' or 3' flanking sequences or intronic sequences naturally associated with the subject genes or naturally occurring mutants thereof. The polynucleotide of a cell is rendered accessible for hybridization, the probe is contacted with the polynucleotide of the sample, and the hybridization of the probe to the sample polynucleotide is detected. Such techniques can be used to detect lesions or allelic variants at either the genomic or mRNA level, including deletions, substitutions, etc., as well as to determine mRNA transcript levels.

A preferred detection method is allele specific hybridization using probes overlapping the mutation or polymorphic site and having about 5, 10, 20, 25, or 30 nucleotides around the mutation or polymorphic region. In a preferred embodiment of the invention, several probes capable of hybridising specifically to allelic variants are attached to a solid phase support, e.g., a "chip". Mutation detection analysis using these chips comprising oligonucleotides, also termed "DNA probe arrays" is described e.g., in Cronin et al. (119). In one embodiment, a chip comprises all the allelic variants of at least one polymorphic region of a gene. The solid phase support is then contacted with a test polynucleotide and hybridization to the specific probes is detected. Accordingly, the identity of numerous allelic variants of one or more genes can be identified in a simple hybridization experiment.

In certain embodiments, detection of the lesion comprises utilizing the probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligase chain reaction (LCR) [Landegran et al., 1988, (120) and Nakazawa et al., 1994 (121)], the latter of which can be particularly useful for detecting point mutations in the gene; Abravaya et al., 1995 ,(122)]. In a merely illustrative embodiment, the method includes the steps of (i) collecting a sample of cells from a patient, (ii) isolating polynucleotide (e.g., genomic, mRNA or both) from the cells of the sample, (iii) contacting the polynucleotide sample with one or more primers which specifically hybridize to a polynucleotide sequence under conditions such that hybridization and amplification of the polynucleotide (if present) occurs, and (iv) detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary

amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication [Guatelli, J.C. et al., 1990, (123)], transcriptional amplification system [Kwoh, D.Y. et al., 1989, (124)], Q-Beta replicase [Lizardi, P.M. et al., 1988 ,(125)], or any other polynucleotide amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of polynucleotide molecules if such molecules are present in very low numbers.

In a preferred embodiment of the subject assay, mutations in, or allelic variants, of a gene from a sample cell are identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis. Moreover; the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In situ hybridization

In one aspect, the method comprises *in situ* hybridization with a probe derived from a given marker polynucleotide, which sequence is selected from any of the polynucleotide sequences of the SEQ ID NO: 1 to 9, or 11 to 19 or 21 to 26 and 53 to 75 or a sequence complementary thereto. The method comprises contacting the labeled hybridization probe with a sample of a given type of tissue from a patient potentially having malignant neoplasia and breast cancer in particular as well as normal tissue from a person with no malignant neoplasia, and determining whether the probe labels tissue of the patient to a degree significantly different (e.g., by at least a factor of two, or at least a factor of five, or at least a factor of twenty, or at least a factor of fifty) than the degree to which normal tissue is labelled.

Polypeptide detection

The subject invention further provides a method of determining whether a cell sample obtained from a subject possesses an abnormal amount of marker polypeptide which comprises (a) obtaining a cell sample from the subject, (b) quantitatively determining the amount of the marker polypeptide in the sample so obtained, and (c) comparing the amount of the marker polypeptide so determined with a known standard, so as to thereby determine whether the cell sample obtained from the subject possesses an abnormal amount of the marker polypeptide. Such marker

polypeptides may be detected by immunohistochemical assays, dot-blot assays, ELISA and th like.

Antibodies

Any type of antibody known in the art can be generated to bind specifically to an epitope of a „BREAST CANCER GENE“ polypeptide. An antibody as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, $F(ab)_2$, and Fv, which are capable of binding an epitope of a „BREAST CANCER GENE“ polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, e.g., at least 15, 25, or 50 amino acids.

An antibody which specifically binds to an epitope of a „BREAST CANCER GENE“ polypeptide can be used therapeutically, as well as in immunochemical assays, such as Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody which specifically binds to the immunogen.

Typically, an antibody which specifically binds to a „BREAST CANCER GENE“ polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to „BREAST CANCER GENE“ polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate a „BREAST CANCER GENE“ polypeptide from solution.

„BREAST CANCER GENE“ polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal antibodies. If desired, a „BREAST CANCER GENE“ polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response. Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (e.g., aluminum hydroxide), and surface active substances (e.g. lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially useful.

Monoclonal antibodies which specifically bind to a „BREAST CANCER GENE“ polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B cell hybridoma technique, and the EBV hybridoma technique [Kohler et al., 1985, (136); Kozbor et al., 1985, (137); Cote et al., 1983, (138) and Cole et al., 1984, (139)].

In addition, techniques developed for the production of chimeric antibodies, the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used [Morrison et al., 1984, (140); Neuberger et al., 1984, (141); Takeda et al., 1985, (142)]. Monoclonal and other antibodies also can be humanized to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in therapy or may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively, humanized antibodies can be produced using recombinant methods, as described in GB2188638B. Antibodies which specifically bind to a „BREAST CANCER GENE“ polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. Patent 5,565,332.

Alternatively, techniques described for the production of single chain antibodies can be adapted using methods known in the art to produce single chain antibodies which specifically bind to „BREAST CANCER GENE“ polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain shuffling from random combinatorial immunoglobulin libraries [Burton, 1991, (143)].

Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template [Thirion et al., 1996, (144)]. Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain antibodies is taught, for example, in Coloma & Morrison, (145). Construction of bivalent, bispecific single-chain antibodies is taught in Mallender & Voss, (146).

A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology [Verhaar et al., 1995, (147); Nicholls et al., 1993, (148)].

Antibodies which specifically bind to „BREAST CANCER GENE" polypeptides also can b produced by inducing in vivo production in the lymphocyte population or by screenin immunoglobulin libraries or panels of highly specific binding reagents as disclosed in th literature [Orlandi et al., 1989, (149) and Winter et al., 1991, (150)].

Other types of antibodies can be constructed and used therapeutically in methods of the invention For example, chimeric antibodies can be constructed as disclosed in WO 93/03151. Bindin proteins which are derived from immunoglobulins and which are multivalent and multispecific such as the antibodies described in WO 94/13804, also can be prepared.

Antibodies according to the invention can be purified by methods well known in the art. Fo example, antibodies can be affinity purified by passage over a column to which a „BREAST CANCER GENE" polypeptide is bound. The bound antibodies can then be eluted from the columr using a buffer with a high salt concentration.

Immunoassays are commonly used to quantify the levels of proteins in cell samples, and many other immunoassay techniques are known in the art. The invention is not limited to a particular assay procedure, and therefore is intended to include both homogeneous and heterogeneous procedures. Exemplary immunoassays which can be conducted according to the invention include fluorescence polarisation immunoassay (FPIA), fluorescence immunoassay (FIA), enzyme immunoassay (EIA), nephelometric inhibition immunoassay (NIA), enzyme linked immunosorbent assay (ELISA), and radioimmunoassay (RIA). An indicator moiety, or label group, can be attached to the subject antibodies and is selected so as to meet the needs of various uses of the method which are often dictated by the availability of assay equipment and compatible immunoassay procedures. General techniques to be used in performing the various immunoassays noted above are known to those of ordinary skill in the art.

In another embodiment, the level of at least one product encoded by any of the polynucleotide sequences of the SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19 or 21 to 26 or 53 to 75 or of at least 2 products encoded by a polynucleotide selected from SEQ ID NO: 1 to 26 and 53 to 75 or a sequence complementary thereto, in a biological fluid (e.g., blood or urine) of a patient may be determined as a way of monitoring the level of expression of the marker polynucleotide sequence in cells of that patient. Such a method would include the steps of obtaining a sample of a biological fluid from the patient, contacting the sample (or proteins from the sample) with an antibody specific for a encoded marker polypeptide, and determining the amount of immune complex formation by the antibody, with the amount of immune complex formation being indicative of the level of the marker encoded product in the sample. This determination is

particularly instructive when compared to the amount of immune complex formation by the same antibody in a control sample taken from a normal individual or in one or more samples previously or subsequently obtained from the same person.

In another embodiment, the method can be used to determine the amount of marker polypeptide present in a cell, which in turn can be correlated with progression of the disorder, e.g., plaque formation. The level of the marker polypeptide can be used predictively to evaluate whether a sample of cells contains cells which are, or are predisposed towards becoming, plaque associated cells. The observation of marker polypeptide level can be utilized in decisions regarding, e.g., the use of more stringent therapies.

As set out above, one aspect of the present invention relates to diagnostic assays for determining, in the context of cells isolated from a patient, if the level of a marker polypeptide is significantly reduced in the sample cells. The term "significantly reduced" refers to a cell phenotype wherein the cell possesses a reduced cellular amount of the marker polypeptide relative to a normal cell of similar tissue origin. For example, a cell may have less than about 50%, 25%, 10%, or 5% of the marker polypeptide that a normal control cell. In particular, the assay evaluates the level of marker polypeptide in the test cells, and, preferably, compares the measured level with marker polypeptide detected in at least one control cell, e.g., a normal cell and/or a transformed cell of known phenotype.

Of particular importance to the subject invention is the ability to quantify the level of marker polypeptide as determined by the number of cells associated with a normal or abnormal marker polypeptide level. The number of cells with a particular marker polypeptide phenotype may then be correlated with patient prognosis. In one embodiment of the invention, the marker polypeptide phenotype of the lesion is determined as a percentage of cells in a biopsy which are found to have abnormally high/low levels of the marker polypeptide. Such expression may be detected by immunohistochemical assays, dot-blot assays, ELISA and the like.

Immunohistochemistry

Where tissue samples are employed, immunohistochemical staining may be used to determine the number of cells having the marker polypeptide phenotype. For such staining, a multiblock of tissue is taken from the biopsy or other tissue sample and subjected to proteolytic hydrolysis, employing such agents as protease K or pepsin. In certain embodiments, it may be desirable to isolate a nuclear fraction from the sample cells and detect the level of the marker polypeptide in the nuclear fraction.

The tissues samples are fixed by treatment with a reagent such as formalin, glutaraldehy methanol, or the like. The samples are then incubated with an antibody, preferably a monoclo antibody, with binding specificity for the marker polypeptides. This antibody may be conjugat to a Label for subsequent detection of binding. samples are incubated for a time Sufficient f formation of the immunocomplexes. Binding of the antibody is then detected by virtue of a Lat conjugated to this antibody. Where the antibody is unlabelled, a second labeled antibody may l employed, e.g., which is specific for the isotype of the anti-marker polypeptide antibod Examples of labels which may be employed include radionuclides, fluorescenc chemiluminescence, and enzymes.

Where enzymes are employed, the Substrate for the enzyme may be added to the samples t provide a colored or fluorescent product. Examples of suitable enzymes for use in conjugate include horseradish peroxidase, alkaline phosphatase, malate dehydrogenase and the like. Wher not commercially available, such antibody-enzyme conjugates are readily produced by technique known to those skilled in the art.

In one embodiment, the assay is performed as a dot blot assay. The dot blot assay finds particula application where tissue samples are employed as it allows determination of the average amount o the marker polypeptide associated with a Single cell by correlating the amount of marke polypeptide in a cell-free extract produced from a predetermined number of cells.

In yet another embodiment, the invention contemplates using one or more antibodies which are generated against one or more of the marker polypeptides of this invention, which polypeptides are encoded by any of the polynucleotide sequences of the SEQ ID NO: 1 to 26 or 53 to 75. Such a panel of antibodies may be used as a reliable diagnostic probe for breast cancer. The assay of the present invention comprises contacting a biopsy sample containing cells, e.g., macrophages, with a panel of antibodies to one or more of the encoded products to determine the presence or absence of the marker polypeptides.

The diagnostic methods of the subject invention may also be employed as follow-up to treatment, e.g., quantification of the level of marker polypeptides may be indicative of the effectiveness of current or previously employed therapies for malignant neoplasia and breast cancer in particular as well as the effect of these therapies upon patient prognosis.

The diagnostic assays described above can be adapted to be used as prognostic assays, as well. Such an application takes advantage of the sensitivity of the assays of the Invention to events which take place at characteristic stages in the progression of plaque generation in case of

malignant neoplasia. For example, a given marker gene may be up- or down-regulated at a very early stage, perhaps before the cell is developing into a foam cell, while another marker gene may be characteristically up or down regulated only at a much later stage. Such a method could involve the steps of contacting the mRNA of a test cell with a polynucleotide probe derived from a given marker polynucleotide which is expressed at different characteristic levels in breast cancer tissue cells at different stages of malignant neoplasia progression, and determining the approximate amount of hybridization of the probe to the mRNA of the cell, such amount being an indication of the level of expression of the gene in the cell, and thus an indication of the stage of disease progression of the cell; alternatively, the assay can be carried out with an antibody specific for the gene product of the given marker polynucleotide, contacted with the proteins of the test cell. A battery of such tests will disclose not only the existence of a certain arteriosclerotic plaque, but also will allow the clinician to select the mode of treatment most appropriate for the disease, and to predict the likelihood of success of that treatment.

The methods of the invention can also be used to follow the clinical course of a given breast cancer predisposition. For example, the assay of the Invention can be applied to a blood sample from a patient; following treatment of the patient for BREAST CANCER, another blood sample is taken and the test repeated. Successful treatment will result in removal of demonstrate differential expression, characteristic of the breast cancer tissue cells, perhaps approaching or even surpassing normal levels.

Polypeptide activity

In one embodiment the present invention provides a method for screening potentially therapeutic agents which modulate the activity of one or more "BREAST CANCER GENE" polypeptides, such that if the activity of the polypeptide is increased as a result of the upregulation of the "BREAST CANCER GENE" in a subject having or at risk for malignant neoplasia and breast cancer in particular, the therapeutic substance will decrease the activity of the polypeptide relative to the activity of the some polypeptide in a subject not having or not at risk for malignant neoplasia or breast cancer in particular but not treated with the therapeutic agent. Likewise, if the activity of the polypeptide as a result of the downregulation of the "BREAST CANCER GENE" is decreased in a subject having or at risk for malignant neoplasia or breast cancer in particular, the therapeutic agent will increase the activity of the polypeptide relative to the activity of the same polypeptide in a subject not having or not at risk for malignant neoplasia or breast cancer in particular, but not treated with the therapeutic agent.

The activity of the "BREAST CANCER GENE" polypeptides indicated in Table 2 or 3 may measured by any means known to those of skill in the art, and which are particular for the type activity performed by the particular polypeptide. Examples of specific assays which may be us to measure the activity of particular polynucleotides are shown below.

a) G protein coupled receptors

In one embodiment, the "BREAST CANCER GENE" polynucleotide may encode a G prote: coupled receptor. In one embodiment, the present invention provides a method of screenin potential modulators (inhibitors or activators) of the G protein coupled receptor by measurin changes in the activity of the receptor in the presence of a candidate modulator.

1. G_i–coupled receptors

Cells (such as CHO cells or primary cells) are stably transfected with the relevant receptor an with an inducible CRE-luciferase construct. Cells are grown in 50% Dulbecco's modified Eagl medium / 50% F12 (DMEM/F12) supplemented with 10% FBS, at 37°C in a humidifie atmosphere with 10% CO_2 and are routinely split at a ratio of 1:10 every 2 or 3 days. Test culture are seeded into 384 – well plates at an appropriate density (e.g. 2000 cells / well in 35 µl cel culture medium) in DMEM/F12 with FBS, and are grown for 48 hours (range: ~ 24 - 60 hours depending on cell line). Growth medium is then exchanged against serum free medium (SFM; e.g. Ultra-CHO), containing 0,1% BSA. Test compounds dissolved in DMSO are diluted in SFM and transferred to the test cultures (maximal final concentration 10 µmolar), followed by addition of forskolin (~ 1 µmolar, final conc.) in SFM + 0,1% BSA 10 minutes later. In case of antagonist screening both, an appropriate concentration of agonist, and forskolin are added. The plates are incubated at 37°C in 10% CO_2 for 3 hours. Then the supernatant is removed, cells are lysed with lysis reagent (25 mmolar phosphate-buffer, pH 7,8, containing 2 mmolar DDT, 10% glycerol and 3% Triton X100). The luciferase reaction is started by addition of substrate-buffer (e.g. luciferase assay reagent, Promega) and luminescence is immediately determined (e.g. Berthold luminometer or Hamamatzu camera system).

2. G_s–coupled receptors

Cells (such as CHO cells or primary cells) are stably transfected with the relevant receptor and with an inducible CRE-luciferase construct. Cells are grown in 50% Dulbecco's modified Eagle medium / 50% F12 (DMEM/F12) supplemented with 10% FBS, at 37°C in a humidified

atmosphere with 10% CO_2 and are routinely split at a ratio of 1:10 every 2 or 3 days. Test cultures are seeded into 384 – well plates at an appropriate density (e.g. 1000 or 2000 cells / well in 35 µl cell culture medium) in DMEM/F12 with FBS, and are grown for 48 hours (range: ~ 24 - 60 hours, depending on cell line). The assay is started by addition of test-compounds in serum free medium (SFM; e.g. Ultra-CHO) containing 0,1% BSA: Test compounds are dissolved in DMSO, diluted in SFM and transferred to the test cultures (maximal final concentration 10 µmolar, DMSO conc. < 0,6 %). In case of antagonist screening an appropriate concentration of agonist is added 5 – 10 minutes later. The plates are incubated at 37°C in 10% CO_2 for 3 hours. Then the cells are lysed with 10 µl lysis reagent per well (25 mmolar phosphate-buffer, pH 7,8 , containing 2 mmolar DDT, 10% glycerol and 3% Triton X100) and the luciferase reaction is started by addition of 20 µl substrate-buffer per well (e.g. luciferase assay reagent, Promega). Measurement of luminescence is started immediately (e.g. Berthold luminometer or Hamamatzu camera system).

3. G_q –coupled receptors

Cells (such as CHO cells or primary cells) are stably transfected with the relevant receptor. Cells expressing functional receptor protein are grown in 50% Dulbecco's modified Eagle medium / 50% F12 (DMEM/F12) supplemented with 10% FBS, at 37°C in a humidified atmosphere with 5% CO_2 and are routinely split at a cell line dependent ratio every 3 or 4 days. Test cultures are seeded into 384 – well plates at an appropriate density (e.g. 2000 cells / well in 35 µl cell culture medium) in DMEM/F12 with FBS, and are grown for 48 hours (range: ~ 24 - 60 hours, depending on cell line). Growth medium is then exchanged against physiological salt solution (e.g. Tyrode solution). Test compounds dissolved in DMSO are diluted in Tyrode solution containing 0.1% BSA and transferred to the test cultures (maximal final concentration 10 µmolar). After addition of the receptor specific agonist the resulting Gq-mediated intracellular calcium increase is measured using appropriate read-out systems (e.g. calcium-sensitive dyes).

b) Ion channels

Ion channels are integral membrane proteins involved in electrical signaling, transmembrane signal transduction, and electrolyte and solute transport. By forming macromolecular pores through the membrane lipid bilayer, ion channels account for the flow of specific ion species driven by the electrochemical potential gradient for the permeating ion. At the single molecule level, individual channels undergo conformational transitions ("gating") between the 'open' (ion conducting) and 'closed' (non conducting) state. Typical single channel openings last for a few milliseconds and result in elementary transmembrane currents in the range of 10^{-9} - 10^{-12} Ampere. Channel gating is controlled by various chemical and/or biophysical parameters, such as

neurotransmitters and intracellular second messengers ('ligand-gated' channels) or membrane potential ('voltage-gated' channels). Ion channels are functionally characterized by their ion selectivity, gating properties, and regulation by hormones and pharmacological agents. Because of their central role in signaling and transport processes, ion channels present ideal targets for pharmacological therapeutics in various pathophysiological settings.

In one embodiment, the "BREAST CANCER GENE" may encode an ion channel. In one embodiment, the present invention provides a method of screening potential activators or inhibitors of channels activity of the "BREAST CANCER GENE" polypeptide. Screening for compounds interaction with ion channels to either inhibit or promote their activity can be based on (1.) binding and (2.) functional assays in living cells[Hille (183)].

1. For ligand-gated channels, e.g. ionotropic neurotransmitter/hormone receptors, assays can be designed detecting binding to the target by competition between the compound and a labeled ligand.

2. Ion channel function can be tested functionally in living cells. Target proteins are either expressed endogenously in appropriate reporter cells or are introduced recombinantly. Channel activity can be monitored by (2.1) concentration changes of the permeating ion (most prominently Ca^{2+} ions), (2.2) by changes in the transmembrane electrical potential gradient, and (2.3) by measuring a cellular response (e.g. expression of a reporter gene, secretion of a neurotransmitter) triggered or modulated by the target activity.

2.1 Channel activity results in transmembrane ion fluxes. Thus activation of ionic channels can be monitored by the resulting changes in intracellular ion concentrations using luminescent or fluorescent indicators. Because of its wide dynamic range and availability of suitable indicators this applies particularly to changes in intracellular Ca^{2+} ion concentration ($[Ca^{2+}]_i$). $[Ca^{2+}]_i$ can be measured, for example, by aequorin luminescence or fluorescence dye technology (e.g. using Fluo-3, Indo-1, Fura-2). Cellular assays can be designed where either the Ca^{2+} flux through the target channel itself is measured directly or where modulation of the target channel affects membrane potential and thereby the activity of co-expressed voltage-gated Ca^{2+} channels.

2.2 Ion channel currents result in changes of electrical membrane potential (V_m) which can be monitored directly using potentiometric fluorescent probes. These electrically charged indicators (e.g. the anionic oxonol dye $DiBAC_4(3)$) redistribute between extra- and

intracellular compartment in response to voltage changes. The equilibrium distribution is governed by the Nernst-equation. Thus changes in membrane potential results in concomitant changes in cellular fluorescence. Again, changes in V_m might be caused directly by the activity of the target ion channel or through amplification and/or prolongation of the signal by channels co-expressed in the same cell.

2.3 Target channel activity can cause cellular Ca^{2+} entry either directly or through activation of additional Ca^{2+} channel (see 2.1). The resulting intracellular Ca^{2+} signals regulate a variety of cellular responses, e.g. secretion or gene transcription. Therefore modulation of the target channel can be detected by monitoring secretion of a known hormone/transmitter from the target-expressing cell or through expression of a reporter gene (e.g. luciferase) controlled by an Ca^{2+}-responsive promoter element (e.g. cyclic AMP/ Ca^{2+}-responsive elements; CRE).

c) DNA-binding proteins and transcription factors

In one embodiment, the "BREAST CANCER GENE" may encode a DNA-binding protein or a transcription factor. The activity of such a DNA-binding protein or a transcription factor may be measured, for example, by a promoter assay which measures the ability of the DNA-binding protein or the transcription factor to initiate transcription of a test sequence linked to a particular promoter. In one embodiment, the present invention provides a method of screening test compounds for its ability to modulate the activity of such a DNA-binding protein or a transcription factor by measuring the changes in the expression of a test gene which is regulated by a promoter which is responsive to the transcription factor.

d) Promotor assays

A promoter assay was set up with a human hepatocellular carcinoma cell HepG2 that was stably transfected with a luciferase gene under the control of a gene of interest (e.g. thyroid hormone) regulated promoter. The vector 2xIROluc, which was used for transfection, carries a thyroid hormone responsive element (TRE) of two 12 bp inverted palindromes separated by an 8 bp spacer in front of a tk minimal promoter and the luciferase gene. Test cultures were seeded in 96 well plates in serum - free Eagle's Minimal Essential Medium supplemented with glutamine, tricine, sodium pyruvate, non – essential amino acids, insulin, selen, transferrin, and were cultivated in a humidified atmosphere at 10 % CO_2 at 37°C. After 48 hours of incubation serial dilutions of test compounds or reference compounds (L-T3, L-T4 e.g.) and co-stimulator if appropriate (final concentration 1 nM) were added to the cell cultures and incubation was continued for the optimal

time (e.g. another 4-72 hours). The cells were then lysed by addition of buffer containing Trit X100 and luciferin and the luminescence of luciferase induced by T3 or other compounds w measured in a luminometer. For each concentration of a test compound replicates of 4 were teste EC_{50} – values for each test compound were calculated by use of the Graph Pad Prism Scientif software.

Screening Methods

The invention provides assays for screening test compounds which bind to or modulate the activit of a „BREAST CANCER GENE" polypeptide or a „BREAST CANCER GENE" polynucleotid A test compound preferably binds to a „BREAST CANCER GENE" polypeptide c polynucleotide. More preferably, a test compound decreases or increases „BREAST CANCE GENE" activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the test compound.

Test Compounds

Test compounds can be pharmacological agents already known in the art or can be compound previously unknown to have any pharmacological activity. The compounds can be naturall occurring or designed in the laboratory. They can be isolated from microorganisms, animals, o plants, and can be produced recombinant, or synthesised by chemical methods known in the art. I desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the one-bead one-compound library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds. [For review see Lam, 1997, (151)].

Methods for the synthesis of molecular libraries are well known in the art [see, for example, DeWitt et al., 1993, (152); Erb et al., 1994, (153); Zuckermann et al., 1994, (154); Cho et al., 1993, (155); Carell et al., 1994, (156) and Gallop et al., 1994, (157). Libraries of compounds can be presented in solution [see, e.g., Houghten, 1992, (158)], or on beads [Lam, 1991, (159)], DNA-chips [Fodor, 1993, (160)], bacteria or spores (Ladner, U.S. Patent 5,223,409), plasmids [Cull et al., 1992, (161)], or phage [Scott & Smith, 1990, (162); Devlin, 1990, (163); Cwirla et al., 1990, (164); Felici, 1991, (165)].

High Throughput Screening

Test compounds can be screened for the ability to bind to „BREAST CANCER GENE" polypeptides or polynucleotides or to affect „BREAST CANCER GENE" activity or „BREAST CANCER GENE" expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well, 384-well or 1536-well microtiter plates. The wells of the microtiter plates typically require assay volumes that range from 5 to 500 µl. In addition to the plates, many instruments, materials, pipettors, robotics, plate washers, and plate readers are commercially available to fit the microwell formats.

Alternatively, free format assays, or assays that have no physical barrier between samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by Jayawickreme et al., (166). The cells are placed under agarose in culture dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds are partially released the compounds from the beads. Active compounds can be visualised as dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, (167). Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

In another example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar [Salmon et al., 1996, (168)].

Another high throughput screening method is described in Beutel et al., U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support. When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

Binding Assays

For binding assays, the test compound is preferably a small molecule which binds to and occupie: for example, the ATP/GTP binding site of the enzyme or the active site of a „BREAST CANCEI GENE" polypeptide, such that normal biological activity is prevented. Examples of such smal molecules include, but are not limited to, small peptides or peptide-like molecules.

In binding assays, either the test compound or a „BREAST CANCER GENE" polypeptide cai comprise a detectable label, such as a fluorescent, radioisotopic, chemiluminescent, or enzymati label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a tes compound which is bound to a „BREAST CANCER GENE" polypeptide can then b accomplished, for example, by direct counting of radioemmission, by scintillation counting, or b determining conversion of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to a „BREAST CANCER GENE" polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with a „BREAST CANCER GENE" polypeptide. A microphysiometer (e.g., CytosensorJ) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and a „BREAST CANCER GENE" polypeptide [McConnell et al., 1992, (169)].

Determining the ability of a test compound to bind to a „BREAST CANCER GENE" polypeptide also can be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA) [Sjolander & Urbaniczky, 1991, (170), and Szabo et al., 1995, (171)]. BIA is a technology for studying biospecific interactions in real time, without labeling any of the interactants (e.g., BIAcore™). Changes in the optical phenomenon surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In yet another aspect of the invention, a „BREAST CANCER GENE" polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay [see, e.g., U.S. Patent 5,283,317; Zervos et al., 1993, (172); Madura et al., 1993, (173); Bartel et al., 1993, (174); Iwabuchi et al., 1993, (175) and Brent WO 94/10300], to identify other proteins which bind to or interact with the „BREAST CANCER GENE" polypcptidc and modulatc its activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct, polynucleotide encoding a „BREAST CANCER GENE"

polypeptide can be fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (e.g., GAL4). In the other construct a DNA sequence that encodes an unidentified protein ("prey" or "sample") can be fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact in vivo to form an protein- dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein which interacts with the „BREAST CANCER GENE“ polypeptide.

It may be desirable to immobilize either a „BREAST CANCER GENE“ polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound forms of one or both of the interactants, as well as to accommodate automation of the assay. Thus, either a „BREAST CANCER GENE“ polypeptide (or polynucleotide) or the test compound can be bound to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be used to attach a „BREAST CANCER GENE“ polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached respectively to the polypeptide (or polynucleotide) or test compound and the solid support. Test compounds are preferably bound to the solid support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to a „BREAST CANCER GENE“ polypeptide (or polynucleotide) can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, a „BREAST CANCER GENE“ polypeptide is a fusion protein comprising a domain that allows the „BREAST CANCER GENE“ polypeptide to be bound to a solid support. For example, glutathione S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the nonadsorbed „BREAST CANCER GENE“ polypeptide; the mixture is then incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the

interactants can be determined either directly or indirectly, as described above. Alternatively, t complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilising proteins or polynucleotides on a solid support also can be us in the screening assays of the invention. For example, either a „BREAST CANCER GEN polypeptide (or polynucleotide) or a test compound can be immobilized utilizing conjugation biotin and streptavidin. Biotinylated „BREAST CANCER GENE" polypeptides (polynucleotides) or test compounds can be prepared from biotin NHS (N-hydroxysuccinimid using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, Ill and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical Alternatively, antibodies which specifically bind to a „BREAST CANCER GENE" polypeptid polynucleotide, or a test compound, but which do not interfere with a desired binding site, such a the ATP/GTP binding site or the active site of the „BREAST CANCER GENE" polypeptide, ca be derivatised to the wells of the plate. Unbound target or protein can be trapped in the wells b antibody conjugation.

Methods for detecting such complexes, in addition to those described above for the GST immobilized complexes, include immunodetection of complexes using antibodies whicl specifically bind to a „BREAST CANCER GENE" polypeptide or test compound, enzyme-linkec assays which rely on detecting an activity of a „BREAST CANCER GENE" polypeptide, and SDS gel electrophoresis under non-reducing conditions.

Screening for test compounds which bind to a „BREAST CANCER GENE" polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises a „BREAST CANCER GENE" polypeptide or polynucleotide can be used in a cell-based assay system. A „BREAST CANCER GENE" polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Binding of the test compound to a „BREAST CANCER GENE" polypeptide or polynucleotide is determined as described above.

Modulation of Gene Expression

In another embodiment, test compounds which increase or decrease „BREAST CANCER GENE" expression are identified. A „BREAST CANCER GENE" polynucleotide is contacted with a test compound, and the expression of an RNA or polypeptide product of the „BREAST CANCER GENE" polynucleotide is determined. The level of expression of appropriate mRNA or polypeptide in the presence of the test compound is compared to the level of expression of mRNA or polypeptide in the absence of the test compound. The test compound can then be identified as a

modulator of expression based on this comparison. For example, when expression of mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of the mRNA or polypeptide expression. Alternatively, when expression of the mRNA or polypeptide is less in the presence of the test compound than in its absence, the test compound is identified as an inhibitor of the mRNA or polypeptide expression.

The level of „BREAST CANCER GENE“ mRNA or polypeptide expression in the cells can be determined by methods well known in the art for detecting mRNA or polypeptide. Either qualitative or quantitative methods can be used. The presence of polypeptide products of a „BREAST CANCER GENE“ polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined in vivo, in a cell culture, or in an in vitro translation system by detecting incorporation of labeled amino acids into a „BREAST CANCER GENE“ polypeptide.

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell which expresses a „BREAST CANCER GENE“ polynucleotide can be used in a cell-based assay system. A „BREAST CANCER GENE“ polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, such as CHO or human embryonic kidney 293 cells, can be used.

Therapeutic Indications and Methods

Therapies for treatment of breast cancer primarily relied upon effective chemotherapeutic drugs for intervention on the cell proliferation, cell growth or angiogenesis. The advent of genomics-driven molecular target identification has opened up the possibility of identifying new breast cancer-specific targets for therapeutic intervention that will provide safer, more effective treatments for malignant neoplasia patients and breast cancer patients in particular. Thus, newly discovered breast cancer-associated genes and their products can be used as tools to develop innovative therapies. The identification of the Her2/neu receptor kinase presents exciting new opportunities for treatment of a certain subset of tumor patients as described before. Genes playing important roles in any of the physiological processes outlined above can be characterized as breast cancer targets. Genes or gene fragments identified through genomics can readily be expressed in one or more heterologous expression systems to produce functional recombinant proteins. These proteins are characterized in vitro for their biochemical properties and then used as tools in high-throughput molecular screening programs to identify chemical modulators of their biochemical

activities. Modulators of target gene expression or protein activity can be identified in this mann and subsequently tested in cellular and in vivo disease models for therapeutic activit Optimization of lead compounds with iterative testing in biological models and detaile pharmacokinetic and toxicological analyses form the basis for drug development and subseque testing in humans.

This invention further pertains to the use of novel agents identified by the screening assay described above. Accordingly, it is within the scope of this invention to use a test compoun identified as described herein in an appropriate animal model. For example, an agent identified a described herein (e.g., a modulating agent, an antisense polynucleotide molecule, a specifi antibody, ribozyme, or a human „BREAST CANCER GENE“ polypeptide binding molecule) ca be used in an animal model to determine the efficacy, toxicity, or side effects of treatment wit such an agent. Alternatively, an agent identified as described herein can be used in an anima model to determine the mechanism of action of such an agent. Furthermore, this invention pertain to uses of novel agents identified by the above described screening assays for treatments a described herein.

A reagent which affects human „BREAST CANCER GENE“ activity can be administered to a human cell, either in vitro or in vivo, to reduce or increase human „BREAST CANCER GENE“ activity. The reagent preferably binds to an expression product of a human „BREAST CANCER GENE“. If the expression product is a protein, the reagent is preferably an antibody. For treatment of human cells ex vivo, an antibody can be added to a preparation of stem cells which have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

In one embodiment, the reagent is delivered using a liposome. Preferably, the liposome is stable in the animal into which it has been administered for at least about 30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung, liver, spleen, heart brain, lymph nodes, and skin.

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about 0.5 µg of DNA per 16 nmol of liposome delivered to about 10^6 cells, more preferably about 1.0 µg of DNA per 16 nmol of liposome delivered to about

10^6 cells, and even more preferably about 2.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes usually used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid composition and/or liposomes having a cholesterol backbone conjugated to polyethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a particular cell type, such as a cell-specific ligand exposed on the outer surface of the liposome.

Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods which are standard in the art (see, for example, U.S. Patent 5,705,151). Preferably, from about 0.1 μg to about 10 μg of polynucleotide is combined with about 8 nmol of liposomes, more preferably from about 0.5 μg to about 5 μg of polynucleotides are combined with about 8 nmol liposomes, and even more preferably about 1.0 μg of polynucleotides is combined with about 8 nmol liposomes.

In another embodiment, antibodies can be delivered to specific tissues in vivo using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, Findeis et al., 1993, (176); Chiou et al., 1994, (177); Wu & Wu, 1988, (178); Wu et al., 1994, (179); Zenke et al., 1990, (180); Wu et al., 1991, (181).

Determination of a Therapeutically Effective Dose

The determination of a therapeutically effective dose is well within the capability of those skilled in the art. A therapeutically effective dose refers to that amount of active ingredient which increases or decreases human „BREAST CANCER GENE" activity relative to the human „BREAST CANCER GENE" activity which occurs in the absence of the therapeutically effective dose.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

Therapeutic efficacy and toxicity, e.g., ED_{50} (the dose therapeutically effective in 50% of the population) and LD_{50} (the dose lethal to 50% of the population), can be determined by standard

pharmaceutical procedures in cell cultures or experimental animals. The dose ratio of toxic t therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD_{50}/ED_{50}.

Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active ingredient or to maintain the desired effect. Factors which can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either ex vivo or in vivo using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads, protoplast fusion, viral infection, electroporation, a gene gun, and DEAE- or calcium phosphate-mediated transfection.

Effective in vivo dosages of an antibody are in the range of about 5 μg to about 50 μg/kg, about 50 μg to about 5 mg/kg, about 100 μg to about 500 μg/kg of patient body weight, and about 200 to about 250 μg/kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective in vivo dosages are in the range of about 100 ng to about 200 ng,

500 ng to about 50 mg, about 1 µg to about 2 mg, about 5 µg to about 500 µg, and about 20 µg to about 100 µg of DNA.

If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides which express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

Preferably, a reagent reduces expression of a „BREAST CANCER GENE“ gene or the activity of a "BREAST CANCER GENE" polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of a „BREAST CANCER GENE“ gene or the activity of a „BREAST CANCER GENE“ polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to „BREAST CANCER GENE“ specific mRNA, quantitative RT-PCR, immunologic detection of a „BREAST CANCER GENE“ polypeptide, or measurement of „BREAST CANCER GENE“ activity.

In any of the embodiments described above, any of the pharmaceutical compositions of the invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy can be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents can act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Any of the therapeutic methods described above can be applied to any subject in need of such therapy, including, for example, birds and mammals such as dogs, cats, cows, pigs, sheep, goats, horses, rabbits, monkeys, and most preferably, humans.

All patents and patent applications cited in this disclosure are expressly incorporated herein by reference. The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

Pharmaceutical Compositions

The invention also provides pharmaceutical compositions which can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise, for example, a „BREAST CANCER GENE“ polypeptide, „BREAST CANCER GENE“ polynucleo-

tide, ribozymes or antisense oligonucleotides, antibodies which specifically bind to a „BREAS CANCER GENE“ polypeptide, or mimetics, agonists, antagonists, or inhibitors of a „BREAS CANCER GENE“ polypeptide activity. The compositions can be administered alone or i combination with at least one other agent, such as stabilizing compound, which can t administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited t(saline, buffered saline, dextrose, and water. The compositions can be administered to a patier alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitabl pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitat processing of the active compounds into preparations which can be used pharmaceutically Pharmaceutical compositions of the invention can be administered by any number of route: including, but not limited to, oral, intravenous, intramuscular, intraarterial, intramedullary intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable foi oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, foi ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethylcellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention can be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 150 mM histidine, 0.1%2% sucrose, and 27% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (182). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of administration.

MATERIAL AND METHODS

One strategy for identifying genes that are involved in breast cancer is to detect genes that ar expressed differentially under conditions associated with the disease versus non-diseas conditions. The sub-sections below describe a number of experimental systems which may be use to detect such differentially expressed genes. In general, these experimental systems include a least one experimental condition in which subjects or samples are treated in a manner associate with breast cancer, in addition to at least one experimental control condition lacking such diseas associated treatment. Differentially expressed genes are detected, as described below, b comparing the pattern of gene expression between the experimental and control conditions.

Once a particular gene has been identified through the use of one such experiment, its expressior pattern may be further characterized by studying its expression in a different experiment and the findings may be validated by an independent technique. Such use of multiple experiments may be useful in distinguishing the roles and relative importance of particular genes in breast cancer. A combined approach, comparing gene expression pattern in cells derived from breast cancer patients to those of *in vitro* cell culture models can give substantial hints on the pathways involved in development and/or progression of breast cancer.

Among the experiments which may be utilized for the identification of differentially expressed genes involved in malignant neoplasia and breast cancer, for example, are experiments designed to analyze those genes which are involved in signal transduction. Such experiments may serve to identify genes involved in the proliferation of cells.

Below are methods described for the identification of genes which are involved in breast cancer. Such represent genes which are differentially expressed in breast cancer conditions relative to their expression in normal, or non-breast cancer conditions or upon experimental manipulation based on clinical observations. Such differentially expressed genes represent "target" and/or "marker" genes. Methods for the further characterization of such differentially expressed genes, and for their identification as target and/or marker genes, are presented below.

Alternatively, a differentially expressed gene may have its expression modulated, i.e., quantitatively increased or decreased, in normal versus breast cancer states, or under control versus experimental conditions. The degree to which expression differs in normal versus breast cancer or control versus experimental states need only be large enough to be visualized via standard characterization techniques, such as, for example, the differential display technique described below. Other such standard characterization techniques by which expression differences

may be visualized include but are not limited to quantitative RT-PCR and Northern analyses, which are well known to those of skill in the art.

As part of this invention, a method is described by way of illustration and not by limitation, displaying at least some of the below mentioned aspects:

1. A method for the prediction, diagnosis or prognosis of malignant neoplasia by the detection of at least 2 markers characterized in that the markers are genes and fragments thereof or genomic nucleic acid sequences that are located on one chromosomal region which is altered in malignant neoplasia.

2. A method for the prediction, diagnosis or prognosis of malignant neoplasia by the detection of at least 2 markers characterized in that the markers are:

    a) genes that are located on one or more chromosomal region(s) which is/are altered in malignant neoplasia; and

    b)

        i) receptor and ligand; or

        ii) members of the same signal transduction pathway; or

        iii) members of synergistic signal transduction pathways; or

        iv) members of antagonistic signal transduction pathways; or

        v) transcription factor and transcription factor binding site.

3. The method of aspect 1 or 2 wherein the malignant neoplasia is breast cancer, ovarian cancer, gastric cancer, colon cancer, esophageal cancer, mesenchymal cancer, bladder cancer or non-small cell lung cancer.

4. The method of aspect 1 or 2 wherein at least one chromosomal region is defined as the cytogenetic region: 1p13, 1q32, 3p21-p24, 5p13-p14, 8q23-q24, 11q13, 12q13,17q12-q24 or 20q13.

5. The method of aspect 1 or 2 wherein at least chromosomal region is defined as the cytogenetic region 17q11.2-21.3 and the malignant neoplasia is breast cancer, ovarian cancer, gastric cancer, colon cancer, esophageal cancer, mesenchymal cancer, bladder cancer or non-small cell lung cancer.

6. The method of aspect 1 or 2 wherein at least one chromosomal region is defined as t cytogenetic region 3p21-24 and the malignant neoplasia is breast cancer, ovarian canc gastric cancer, colon cancer, esophageal cancer, mesenchymal cancer, bladder cancer non-small cell lung cancer.

7. The method of aspect 1 or 2 wherein at least one chromosomal region is defined as tl cytogenetic region 12q13 and the malignant neoplasia is breast cancer, ovarian cance gastric cancer, colon cancer, esophageal cancer, mesenchymal cancer, bladder cancer (non-small cell lung cancer.

8. A method for the prediction, diagnosis or prognosis of malignant neoplasia by th detection of at least one marker whereby the marker is a VNTR, SNP, RFLP or ST characterized in that the marker is located on one chromosomal region which is altered i malignant neoplasia due to amplification and the marker is detected in a cancerous and non-cancerous tissue or biological sample of the same individual.

9. The method of aspect 8 wherein the marker is selected from the group consisting of th VNTRs:

D17S946, D17S1181, D17S2026, D17S838, D17S250, D17S1818, D17S614, D17S2019 D17S608, D17S1655, D17S2147, D17S754, D17S1814, D17S2007, D17S1246, D17S1979, D17S1984, D17S1984, D17S1867, D17S1788, D17S1836, D17S1787, D17S1660, D17S2154, D17S1955, D17S2098, D17S518, D17S1851, D11S4358, D17S964, D19S1091, D17S1179, D10S2160, D17S1230, D17S1338, D17S2011, D17S1237, D17S2038, D17S2091, D17S649, D17S1190 and M87506.

10. The method of aspect 8 wherein the marker is selected from the group consisting of the SNPs:

rs2230698, rs2230700, rs1058808, rs1801200, rs903506, rs2313170, rs1136201, rs2934968, rs2172826, rs1810132, rs1801201, rs2230702, rs2230701, rs1126503, rs3471, rs13695, rs471692, rs558068, rs1064288, rs1061692, rs520630, rs782774, rs565121, rs2586112, rs532299, rs2732786, rs1804539, rs1804538, rs1804537, rs1141364, rs12231, rs1132259, rs1132257, rs1132256, rs1132255, rs1132254, rs1132252, rs1132268 and rs1132258

11. A method for the prediction, diagnosis or prognosis of malignant neoplasia by the detection of at least one marker characterized in that the marker is selected from:

a) a polynucleotide or polynucleotide analog comprising at least one of the sequences of SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26, 53 to 75, or 315 to 318 ;

b) a polynucleotide or polynucleotide analog which hybridizes under stringent conditions to a polynucleotide specified in (a) and encodes a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

c) a polynucleotide or polynucleotide analog the sequence of which deviates from the polynucleotide specified in (a) and (c) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

d) a polynucleotide or polynucleotide analog which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d)

e) a purified polypeptide encoded by a polynucleotide or polynucleotide analog sequence specified in (a) to (e)

f) A purified polypeptide comprising at least one of the sequences of SEQ ID NO: 28 to 32, 34, 35, 37 to 42, 44, 45, 47 to 52, 76 to 98, or 393 to 396;

are detected.

12. A method for the prediction, diagnosis or prognosis of malignant neoplasia by the detection of at least 2 markers characterized in that at least 2 markers are selected from:

a) polynucleotide or polynucleotide analog comprising at least one of the sequences of SEQ ID NO: 1 to 26 or 53 to 75 or 315 to 318;

b) a polynucleotide or polynucleotide analog which hybridizes under stringent conditions to a polynucleotide specified in (a) and encodes a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

c) a polynucleotide or polynucleotide analog the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

d) a polynucleotide or polynucleotide analog which represents a specific fragme derivative or allelic variation of a polynucleotide sequence specified in (a) to (c)

e) a purified polypeptide encoded by a polynucleotide sequence or polynucleoti analog specified in (a) to (d)

f) a purified polypeptide comprising at least one of the sequences of SEQ ID NO: : to 52 or 76 to 98 or 393 to 396

are detected.

13. The method of any of the aspects 1 or 12 wherein the detection method comprises the us of PCR, arrays or beads.

14. A diagnostic kit comprising instructions for conducting the method of any of aspects 1 t 13.

15. A composition for the prediction, diagnosis or prognosis of malignant neoplasi comprising:

a) a detection agent for:

i) any polynucleotide or polynucleotide analog comprising at least one of th sequences of SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26, 53 to 75 or 315 to 318,

ii) any polynucleotide or polynucleotide analog which hybridizes unde stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

iii) a polynucleotide or polynucleotide analog the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

iv) a polynucleotide or polynucleotide analog which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c)

v) a polypeptide encoded by a polynucleotide or polynucleotide analog sequence specified in (a) to (d);

vi) a polypeptide comprising at least one of the sequences of SEQ ID NO: 28 to 32, 34, 35, 37 to 42, 44, 45, 47 to 52, 76 to 98, or 393 to 396.

or

b) at least 2 detection agents for at least 2 markers selected from:

i) any polynucleotide comprising at least one of the sequences of SEQ ID NO: 1 to 26 or 53 to 75 or 315 to 318;

ii) any polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

iii) a polynucleotide the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

iv) a polynucleotide which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c)

v) a polypeptide encoded by a polynucleotide sequence specified in (a) to (d);

vi) a polypeptide comprising at least one of the sequences of SEQ ID NO: [illegible] to 52 or 76 to 98 or 393 to 396.

16. An array comprising a plurality of polynucleotides or polynucleotide analogs wherein each of the polynucleotides is selected from:

a) a polynucleotide or polynucleotide analog comprising at least one of the sequences of SEQ ID NO: 1 to 26 or 53 to 75 or 315 to 318;

b) a polynucleotide or polynucleotide analog which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

c) a polynucleotide or polynucleotide analog the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

d) a polynucleotide or polynucleotide analog which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c)

attached to a solid support.

17. A method of screening for agents which regulate the activity of a polypeptide encoded by a polynucleotide or polynucleotide analog selected from the group consisting of:

a) a polynucleotide or polynucleotide analog comprising at least one of the sequences of SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26, 53 to 75 or 315 to 318;

b) a polynucleotide or polynucleotide analog which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

c) a polynucleotide or polynucleotide analog the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

d) a polynucleotide or polynucleotide analog which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

comprising the steps of:

i) contacting a test compound with at least one polypeptide encoded by a polynucleotide specified in (a) to (d); and

ii) detecting binding of the test compound to the polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential therapeutic agent for modulating the activity of the polypeptide in order to prevent of treat malignant neoplasia.

18. A method of screening for agents which regulate the activity of a polypeptide encoded by a polynucleotide or polynucleotide analog selected from the group consisting of:

    a) a polynucleotide or polynucleotide analog comprising at least one of the sequences of SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26, 53 to 75, or 315 to 318;

    b) a polynucleotide or polynucleotide analog which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

    c) a polynucleotide or polynucleotide analog the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

    d) a polynucleotide or polynucleotide analog which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c)

    comprising the steps of:

    i) contacting a test compound with at least one polypeptide encoded by a polynucleotide specified in (a) to (d); and

    ii) detecting the activity of the polypeptide as specified for the respective sequence in Table 2 or 3, wherein a test compound which increases the activity is identified as a potential preventive or therapeutic agent for increasing the polypeptide aciti[illegible] in malignant neoplasia, and wherein a test compound which decreases the activity of the polypeptide is identified as a potential therapeutic agent for decreasing the polypeptide activity in malignant neoplasia.

19. A method of screening for agents which regulate the activity of a polynucleotide or polynucleotide analog selected from group consisting of;

    a) a polynucleotide or polynucleotide analog comprising at least one of the sequences of SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26, 53 to 75 or 315 to 318;

    b) a polynucleotide or polynucleotide analog which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a polypeptide exhibiting

the same biological function as specified for the respective sequence in Table 2 c 3

c) a polynucleotide or polynucleotide analog the sequence of which deviates from th polynucleotide specified in (a) and (b) due to the generation of the genetic cod encoding a polypeptide exhibiting the same biological function as specified for th respective sequence in Table 2 or 3

d) a polynucleotide or polynucleotide analog which represents a specific fragment derivative or allelic variation of a polynucleotide sequence specified in (a) to (c)

comprising the steps of:

i) contacting a test compound with at least one polynucleotide or polynucleotide analog specified in (a) to (d), and

ii) detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential preventive or therapeutic agent for regulating the activity of the polynucleotide in malignant neoplasia.

20. Use of

a) a polynucleotide or polynucleotide analog comprising at least one of the sequences of SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26, 53 to 75 or 315 to 318;

b) a polynucleotide which hybridizes under stringent conditions to a polynucleotide or polynucleotide analog specified in (a) encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3;

c) a polynucleotide or polynucleotide analog the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3;

d) a polynucleotide or polynucleotide analog which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c);

e) an antisense molecule targeting specifically one of the polynucleotide sequences specified in (a) to (d);

f) a purified polypeptide encoded by a polynucleotide or polynucleotide analog sequence specified in (a) to (d)

g) a purified polypeptide comprising at least one of the sequences of SEQ ID NO: 28 to 32, 34, 35, 37 to 42, 44, 45, 47 to 52, 76 to 98 or 393 to 396;

h) an antibody capable of binding to one of the polynucleotide specified in (a) to (d) or a polypeptide specified in (f) and (g);

i) a reagent identified by any of the methods of aspect 17 to 19 that modulates the amount or activity of a polynucleotide sequence specified in (a) to (d) or a polypeptide specified in (f) and (g);

in the preparation of a composition for the prevention, prediction, diagnosis, prognosis or medicament for the treatment of malignant neoplasia.

21. Use of aspect 20 wherein the disease is breast cancer.

22. A reagent that regulates the activity of a polypeptide selected from the group consisting of:

a) a polypeptide encoded by any polynucleotide or polynucleotide analog comprising at least one of the sequences of SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26, 53 to 75 or 315 to 318;

b) a polypeptide encoded by any polynucleotide or polynucleotide analog which hybridizes under stringent conditions to any polynucleotide comprising at least one of the sequences of SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26, 53 to 75 or 315 to 318 encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

c) a polypeptide encoded by any polynucleotide or polynucleotide analog the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

d) a polypeptide encoded by any polynucleotide or polynucleotide analog which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c)_encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

e) or a polypeptide comprising at least one of the sequences of SEQ ID NO: 28 to 3 34, 35, 37 to 42, 44, 45, 47 to 52, 76 to 98 or 393 to 396;

wherein said reagent is identified by the method of any of the aspects 17 to 19.

23. A reagent that regulates the activity of a polynucleotide or polynucleotide analog selecte from the group consisting of:

a) a polynucleotide or polynucleotide analog comprising at least one of the sequence SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26, 53 to 75 or 315 to 318;

b) a polynucleotide or polynucleotide analog which hybridizes under stringen conditions to a polynucleotide specified in (a) encoding a polypeptide exhibitin the same biological function as specified for the respective sequence in Table 2 o 3

c) a polynucleotide or polynucleotide analog the sequence of which deviates from th polynucleotide specified in (a) and (b) due to the generation of the genetic cod encoding a polypeptide exhibiting the same biological function as specified for th respective sequence in Table 2 or 3

d) a polynucleotide or polynucleotide analog which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c) encoding a polypeptideexhibiting the same biological function as specified for the respective sequence in Table 2 or 3

wherein said reagent is identified by the method of any of the aspects 17 to 19.

24. A pharmaceutical composition, comprising:

a) an expression vector containing at least one polynucleotide or polynucleotide analog selected from the group consisting of:

i) a polynucleotide or polynucleotide analog comprising at least one of the sequences of SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26, 53 to 75 or 315 to 318;

ii) a polynucleotide or polynucleotide analog which hybridizes under stringent conditions to a polynucleotide specified in (a) encoding a

polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

iii) a polynucleotide or polynucleotide analog the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code_encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

iv) a polynucleotide or polynucleotide analog which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (c)_encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3;

or the reagent of aspect 22 or 23 and a pharmaceutically acceptable carrier.

25. A computer-readable medium comprising:

a) at least one digitally encoded value representing a level of expression of at least one polynucleotide sequence of SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26, 53 to 75 or 315 to 318

b) al least 2 digitally encoded values representing the levels of expression of at least 2 polynucleotide sequences selected from SEQ ID NO: 1 to 26, 53 to 75 or 315 to 318

in a cell from the a subject at risk for or having malignant neoplasia.

26. A method for the detection of chromosomal alterations characterized in that the relative abundance of individual mRNAs, encoded by genes, located in altered chromosomal regions is detected.

27. A method for the detection of chromosomal alterations characterized in that the copy number of one or more chromosomal region(s) is detected by quantitative PCR.

EXAMPLE 1

Expression profiling

a) Expression profiling utilizing quantitative RT-PCR

For a detailed analysis of gene expression by quantitative PCR methods, one will utilize prim flanking the genomic region of interest and a fluorescent labeled probe hybridizing in-betwe Using the PRISM 7700 Sequence Detection System of PE Applied Biosystems (Perkin Elm Foster City, CA, USA) with the technique of a fluorogenic probe, consisting of an oligonucleoti labeled with both a fluorescent reporter dye and a quencher dye, one can perform such expression measurement. Amplification of the probe-specific product causes cleavage of tl probe, generating an increase in reporter fluorescence. Primers and probes were selected using tl Primer Express software and localized mostly in the 3' region of the coding sequence or in the untranslated region (see Table 5 for primer- and probe- sequences) according to the relativ positions of the probe sequence used for the construction of the Affymetrix HG_U95A-E or HG U133A-B DNA-chips. All primer pairs were checked for specificity by conventional PC reactions. To standardize the amount of sample RNA, GAPDH was selected as a reference, since was not differentially regulated in the samples analyzed. TaqMan validation experiments wer performed showing that the efficiencies of the target and the control amplifications ar approximately equal which is a prerequisite for the relative quantification of gene expression by the comparative $\Delta\Delta C_T$ method, known to those with skills in the art.

As well as the technology provided by Perkin Elmer one may use other technique implementations like Lightcycler ™ from Roche Inc. or iCycler from Stratagene Inc..

b) Expression profiling utilizing DNA microarrays

Expression profiling can bee carried out using the Affymetrix Array Technology. By hybridization of mRNA to such a DNA-array or DNA-Chip, it is possible to identify the expression value of each transcripts due to signal intensity at certain position of the array. Usually these DNA-arrays are produced by spotting of cDNA, oligonucleotides or subcloned DNA fragments. In case of Affymetrix technology app. 400.000 individual oligonucleotide sequences were synthesized on the surface of a silicon wafer at distinct positions. The minimal length of oligomers is 12 nucleotides, preferable 25 nucleotides or full length of the questioned transcript. Expression profiling may also be carried out by hybridization to nylon or nitro-cellulose membrane bound DNA or oligonucleotides. Detection of signals derived from hybridization may be obtained by either colorimetric, fluorescent, electrochemical, electronic, optic or by radioactive readout. Detailed

description of array construction have been mentioned above and in other patents cited. To determine the quantitative and qualitative changes in the chromosomal region to analyze, RNA from tumor tissue which is suspected to contain such genomic alterations has to be compared to RNA extracted from benign tissue (e.g. epithelial breast tissue, or micro dissected ductal tissue) on the basis of expression profiles for the whole transcriptome. With minor modifications, the sample preparation protocol followed the Affymetrix GeneChip Expression Analysis Manual (Santa Clara, CA). Total RNA extraction and isolation from tumor or benign tissues, biopsies, cell isolates or cell containing body fluids can be performed by using TRIzol (Life Technologies, Rockville, MD) and Oligotex mRNA Midi kit (Qiagen, Hilden, Germany), and an ethanol precipitation step should be carried out to bring the concentration to 1 mg/ml. Using 5–10 mg of mRNA to create double stranded cDNA by the SuperScript system (Life Technologies). First strand cDNA synthesis was primed with a T7-(dT24) oligonucleotide. The cDNA can be extracted with phenol/chloroform and precipitated with ethanol to a final concentration of 1mg /ml. From the generated cDNA, cRNA can be synthesized using Enzo's (Enzo Diagnostics Inc., Farmingdale, NY) *in vitro* Transcription Kit. Within the same step the cRNA can be labeled with biotin nucleotides Bio-11-CTP and Bio-16-UTP (Enzo Diagnostics Inc., Farmingdale, NY) . After labeling and cleanup (Qiagen, Hilden (Germany) the cRNA then should be fragmented in an appropriated fragmentation buffer (e.g., 40 mM Tris-Acetate, pH 8.1, 100 mM KOAc, 30 mM MgOAc, for 35 minutes at 94°C). As per the Affymetrix protocol, fragmented cRNA should be hybridized on the HG_U133 arrays A and B, comprising app. 40.000 probed transcripts each, for 24 hours at 60 rpm in a 45°C hybridization oven. After Hybridization step the chip surfaces have to be washed and stained with streptavidin phycoerythrin (SAPE; Molecular Probes, Eugene, OR) in Affymetrix fluidics stations. To amplify staining, a second labeling step can be introduced, which is recommended but not compulsive. Here one should add SAPE solution twice with an antistreptavidin biotinylated antibody. Hybridization to the probe arrays may be detected by fluorometric scanning (Hewlett Packard Gene Array Scanner; Hewlett Packard Corporation, Palo Alto, CA).

After hybridization and scanning, the microarray images can be analyzed for quality control, looking for major chip defects or abnormalities in hybridization signal. Therefor either Affymetrix GeneChip MAS 5.0 Software or other microarray image analysis software can be utilized. Primary data analysis should be carried out by software provided by the manufacturer..

In case of the genes analyses in one embodiment of this invention the primary data have been analyzed by further bioinformatic tools and additional filter criteria. The bioinformatic analysis is described in detail below.

c) Data analysis

According to Affymetrix measurement technique (Affymetrix GeneChip Expression Analy: Manual, Santa Clara, CA) a single gene expression measurement on one chip yields the avera difference value and the absolute call. Each chip contains 16–20 oligonucleotide probe pairs p gene or cDNA clone. These probe pairs include perfectly matched sets and mismatched sets, bo of which are necessary for the calculation of the average difference, or expression value, measure of the intensity difference for each probe pair, calculated by subtracting the intensity the mismatch from the intensity of the perfect match. This takes into consideration variability hybridization among probe pairs and other hybridization artifacts that could affect the fluorescenc intensities. The average difference is a numeric value supposed to represent the expression valu of that gene. The absolute call can take the values 'A' (absent), 'M' (marginal), or 'P' (presen and denotes the quality of a single hybridization. We used both the quantitative information give by the average difference and the qualitative information given by the absolute call to identify th genes which are differentially expressed in biological samples from individuals with breast cance versus biological samples from the normal population. With other algorithms than the Affymetri: one we have obtained different numerical values representing the same expression values an expression differences upon comparison.

The differential expression E in one of the breast cancer groups compared to the norma population is calculated as follows. Given n average difference values $d_1, d_2, ..., d_n$ in the breas cancer population and m average difference values $c_1, c_2, ..., c_m$ in the population of normal individuals, it is computed by the equation:

$$E \equiv \exp\left(\frac{1}{m}\sum_{i=1}^{m}\ln(c_i) - \frac{1}{n}\sum_{i=1}^{n}\ln(d_i)\right)$$

If d_j<50 or c_i<50 for one or more values of i and j, these particular values c_i and/or d_j are set to an "artificial" expression value of 50. These particular computation of E allows for a correct comparison to TaqMan results.

A gene is called up-regulated in breast cancer versus normal if E≥1.5 and if the number of absolute calls equal to 'P' in the breast cancer population is greater than n/2.

A gene is called down-regulated in breast cancer versus normal if E≤1.5 and if the number of absolute calls equal to 'P' in the normal population is greater than m/2.

The final list of differentially regulated genes consists of all up-regulated and all down-regulated genes in biological samples from individuals with breast cancer versus biological samples from the normal population. Those genes on this list which are interesting for a pharmaceutical application were finally validated by TaqMan. If a good correlation between the expression values/behavior of a transcript could be observed with both techniques, such a gene is listed in Tables 1 to 3.

Since not only the information on differential expression of a single gene within an identified ARCHEON, but also the information on the co-regulation of several members is important for predictive, diagnostic, preventive and therapeutic purposes we have combined expression data with information on the chromosomal position (e.g. golden path) taken from public available databases to develop a picture of the overall transcriptom of a given tumor sample. By this technique not only known or suspected regions of genomes can be inspected but even mo valuable, new regions of disregulation with chromosomal linkage can be identified. This is of value in other types of neoplasia or viral integration and chromosomal rearrangements. By SQL based database searches one can retrieve information on expression, qualitative value of a measurement (denoted by Affymetrix MAS 5.0 Software), expression values derived from other techniques than DNA-chip hybridization and chromosomal linkage.

EXAMPLE 2

Identification of the ARCHEON

a) *Identification and localization of genes or gene probes (represented by the so called probe sets on Affymetrix arrays HG-U95A-E or HG-U133A-B) in their chromosomal context and order on the human genome.*

For identification of larger chromosomal changes or aberrations, as they have been described in detail above, a sufficient number of genes, transcripts or DNA-fragments is needed. The density of probes covering a chromosomal region is not necessarily limited to the transcribed genes, in case of the use of array based CGH but by utilizing RNA as probe material the density is given by the distance of genes on a chromosome. The DNA-microarrays provided by Affymetrix Inc. Do contain hitherto all transcripts from the known humane genome, which are be represented by 40.000 – 60.000 probe sets. By BLAST mapping and sorting the sequences of these short DNA-oligomers to the public available sequence of the human genome represented by the so called "golden path", available at the university of California in Santa Cruz or from the NCBI, a chromosomal display of the whole Transcriptome of a tissue specimen evolves. By graphical display of the individual chromosomal regions and color coding of over or under represented

transcripts, compared to a reference transcriptome regions with DNA gains and losses can b identified.

b) Quantification of gene copy numbers by combined IHC and quantitative PCR (PC karyotyping) or directly by quantitative PCR

Usually one to three paraffin-embedded tissue sections that are 5 µm thick are used to obtai genomic DNA from the samples. Tissue section are stained by colorimetric IHC afte deparaffinization to identify regions containing disease associated cells. Stained regions ar macrodissected with a scalpel and transferred into a micro-centrifuge tube. The genomic DNA o these isolated tissue sections is extracted using appropriate buffers. The isolated DNA is then use for quantitative PCR with appropriate primers and probes. Optionally the IHC staining can be omitted and the genomic DNA can be directly isolated with or without prior deparaffinization with appropriate buffers. Those who are skilled in the art may vary the conditions and buffers described below to obtain equivalent results.

Reagents from DAKO (HercepTest Code No. K 5204) and TaKaRa were used (Biomedicals Cat.: 9091) according to the manufactures protocol.

It is convenient to prepare the following reagents prior to staining:

Solution No. 7

Epitope Retrieval Solution (Citrate buffer + antimicrobial agent) (10xconc.)

20 ml ad 200 ml aqua dest. (stable for 1month at 2-8°C)

Solution No. 8

Washing-buffer (Tris-HCl + antimicrobial agent) (10 x conc.)

30 ml ad 300 ml destilled water (stable for 1month at 2-8°C)

Staining solution: DAB

1 ml solution is sufficient for 10 slides. The solution were prepared immediately before usage.:

1 ml DAB buffer (Substrate Buffer solution, pH 7.5, containing H_2O_2, stabilizer, enhancers and an antimicrobial agent) + 1 drop (25-3 µl) DAB-Chromogen (3,3`-diaminobenzidine chromogen solution). This solution is stable for up to 5 days at 2-8°C. Precipitated substances do not influence the staining result. Additionally required are:2 x approx. 100 ml Xylol, 2 x approx. 100 ml Ethanol

100%, 2 x Ethanol 95%, aqua dest. These solution can be used for up to 40 stainings. A water bath is required for the epitope retrieval step.

Staining procedure:

All reagents are pre-warmed to room temperature (20-25°C) prior to immunostaining. Likewise all incubations were performed at room temperature. Except the epitope retrieval which is performed in at 95°C water bath. Between the steps excess of liquid is tapped off from the slides with lintless tissue (Kim Wipe).

Deparaffinization

Slides are placed in a xylene bath and incubated for 5 minutes. The bath is changed and the step repeated once. Excess of liquid is tapped off and the slides are placed in absolute ethanol for minutes. The bath is changed and the step repeated once. Excess of liquid is tapped off and the slides are placed in 95% ethanol for 3 minutes. The bath is changed and the step repeated once. Excess of liquid is tapped off and the slides are placed in distilled water for a minimum of 30 seconds.

Epitope Retrival

Staining jars are filled with with diluted epitope retrieval solution and preheated in a water bath at 95°C. The deparaffinized sections are immersed into the preheated solution in the staining jars and incubated for 40 minutes at 95°C. The entire jar is removed from the water bath and allowed to cool down at room temperature for 20 minutes. The epitope retrieval solution is decanted, the sections are rinsed in distilled water and finally soaked in wash buffer for 5 minutes.

Peroxidase Blocking:

Excess of buffer is tapped off and the tissue section encircled with a DAKO pen. The specimen is covered with 3 drops (100 µl) Peroxidase-Blocking solution and incubated for 5 minutes. The slides are rinsed in distilled water and placed into a fresh washing buffer bath.

Antibody Incubation

Excess of liquid is tapped off and the specimen are covered with 3 drops (100 µl) of Anti-Her-2/neu reagent (Rabbit Anti-Human Her2 Protein in 0.05 mol/L Tris/HCl, 0.1 mol/L NaCl, 15 mmol/L pH7.2 NaN_3 containing stabilizing protein) or negative control reagent (= IGG fraction

of normal rabbit serum at an equivalent protein concentration as the Her2 Ab). After 30 minutes c incubation the slide is rinsed in water and placed into a fresh water bath.

Visualization

Excess of liquid is tapped off and the specimen are covered with 3 drops (100 µl) of visualizatio reagent. After 30 minutes of incubation the slide is rinsed in water and placed into a fresh wate bath. Excess of liquid is tapped off and the specimen are covered with 3 drops (100 µl) o Substrate-Chromogen solution (DAB) for 10 minutes. After rinsing the specimen with distille water, photographs are taken with a conventional Olympus microscope to document the stainin intensity and tumor regions within the specimen. Optionally a counterstain with hematoxylin wa performed.

DNA extraction

The whole specimens or dissected subregions are transferred into a microcentrifuge tubes. Optionally a small amount (10µl) of preheated TaKaRa solution (DEXPAT™) is preheated and placed onto the specimen to facilitate sample transfer with a scalpel. 50 to 150 µl of TaKaRa solution were added to the samples depending on the size of the tissue sample selected. The sample are incubated at 100°C for 10 minutes in a block heater, followed by centrifugation at 12.000 rpm in a microcentrifuge. The supernatant is collected using a micropet and placed in a separate microcentrifuge tube. If no deparaffinization step has been undertaken one has to be sure not to withdraw tissue debris and resin. Genomic DNA left in the pellet can be collected by adding resin-free TaKaRa buffer and an additional heating and centrifugation step. Samples are stored at -20°C.

Genomic DNA from different tumor cell lines (MCF-7, BT-20, BT-474, SKBR-3, AU-565, UACC-812, UACC-893, HCC-1008, HCC-2157, HCC-1954, HCC-2218, HCC-1937, HCC1599, SW480), or from lymphocytes is prepared with the QIAamp® DNA Mini Kits or the QIAamp® DNA Blood Mini Kits according to the manufacturers protocol. Usually between 1ng up to 1µg DNA is used per reaction.

Quantitative PCR

To measure the gene copy number of the genes within the patient samples the respective primer/probes (see table below) are prepared by mixing 25 µl of the 100 µM stock solution "Upper Primer", 25 µl of the 100 µM stock solution "Lower Primer" with 12,5 µl of the 100 µM stock solution Taq Man Probe (Quencher Tamra) and adjusted to 500 µl with aqua dest. For each

reaction 1,25 µl DNA-Extract of the patient samples or 1,25 µl DNA from the cell lines were mixed with 8,75 µl nuclease-free water and added to one well of a 96 Well-Optical Reaction Plate (Applied Biosystems Part No. 4306737). 1,5 µl Primer/Probe mix, 12, µl Taq Man Universal-PCR Mix (2x) (Applied Biosystems Part No. 4318157) and 1 µl Water are then added. The 96 well plates are closed with 8 Caps/Strips (Applied Biosystems Part Number 4323032) and centrifuged for 3 minutes. Measurements of the PCR reaction are done according to the instructions of the manufacturer with a TaqMan 7900 HT from Applied Biosystems (No. 20114) under appropriate conditions (2 min. 50°C, 10 min. 95°C, 0.15min. 95°C, 1 min. 60°C; 40 cycles). SoftwareSDS 2.0 from Applied Biosysrtems is used according to the respective instructions. CT-values are then further analyzed with appropriate software (Microsoft Excel™).

EXAMPLE 3

Clinical Samples of patients being treated with Herceptin and a chemotherapeutic agent (e.g. docetaxel, paclitaxel, taxotere, carboplatin, cisplatin, oxaliplatin, vinorelbine) as a second line therapy have been obtained. These samples included formalin-fixed and paraffin-embedded material from primary tumours and metastatic lesions of the respective patients. However, the determination of the ARCHEON genes as disclosed in this invention, has also been performed from fresh tissue after nucleic acid extraction in an independent, neoadjuvant setting. Moreover, whole blood, serum and plasma samples were available for multiple patients.

Multiparametric, clinical assessment of the response to Herceptin in combination with chemotherapeutics (e.g. docetaxel, taxotere, paclitaxel, vinorelbine, carboplatin, cisplatin), or other therapies described below, was performed. Clinical information included histological parameters (TNM-Stage, AJCC grade), standard molecular markers (IHC staining for estrogen receptor, progesteron receptor, Her-2/neu) and sonographical or radiological assessment (e.g. CT). Response to treatment was evaluated according to international standards, i.e. modified WHO criteria and RECIST criteria. Each cancer evaluation in the course of the disease was documented (method and date of evaluation, organ, anatomical description, measurability, size of lesion (longest diameter), greatest perpendicular diameter, tumor area). Moreover, each systemic anticancer therapy including prior chemotherapy with anthracyclins (Doxorubicin or Epirubicin) and/or CMF and the response thereto was evaluated (drug, intent, duration, schedule, number of cycles, cumulative dose). The response to combinatory treatment of metatstatic breast cancer patients with Herceptin and chemotherapeutica as second line treatment the modified WHO criteria were used. In addition the initial disease free survival, duration of response and time to progreesion were taken into consideration. For definition of treatment response standard criteria were used: „Complete Response" („CR" = tumor shrinkage of 100 % with no residual disease

being clinical detectable), „Partial Response" („PR" = tumor shrinkage of target lesion of at le 50%), „Stable Disease" („SD" = tumor shrinkage of less than 50 % or no change) a „Progressive Disease" („PD" = tumor growth or new tumor lesions).

More than 70 genes were analyzed according to the method disclosed in example 2 by combir IHC and quantitative PCR or directly by quantitative PCR after nucleic acid extraction from formaldehyde-fixed, paraffin-embedded tissue slides. Results were reconfirmed by independ methodology (VNTR and SNP detection). Alterations of the 43 ARCHEON genes we determined by comparison with reference genes, that are located on the same chromosome intrachromosomal control,) or different chromosomes (= extrachromosomal contro Intrachromosomal reference genes included MMP28, hKa3 and K20. Extrachromosomal referen genes included GAPDH for chromosome 12. However any other gene not included in tl ARCHEONs disclosed in this invention can be used as reference gene for ARCHEO characterizetion. The reference genes should be independent from the ARCHEON alteratio occuring in the neoplastic lesions and should be not affected by chromosomal alterations such a amplifications and deletions. As gene copy numbers of non-amplified genes can be increased i neoplastic lesions due to genomic imbalances such as aneuploidie or polyploidie, eac measurement of ARCHEON genes was correlated to multiple reference genes to minimize th influence of genomic imbalances on the relative copy number calculation. Moreover, minc systemic errors occuring due to differences in the performance of individual primer/probe pair were minimized by determining primer/probe performances in control tissues (i.e. non-neoplasti tissues from healthy controls) and euploid control cell lines (e.g. HS68, ATCC #CRL1635) Moreover one well charcterized, control cell line was used, that displays aneuploidie for a singl chromosome (i.e. Detroit, ATCC#CCL-54; trisomie 21). By measuring genes located on the X chromosome (e.g. SRY), the Y-chromosome (e.g. Xist) and on chromosome 21, defined cop numbers of 1, 2 and 3 genes could be determined as internal control during each run fo standardization. In addition, synthetic targets were spiked into some reactions, that consisted of the target region of the PCR forward and reverse primers of the gene to be normalized, but in between consisted of a synthetic probe hybridization region different from the original probe region of the target gene to be normalized. This allowed internal standardization of each individual qPCR reaction by multiplex PCR. The calculated performance differences were used as a filter for the measurements within the target tissues, i.e. primer/probe differences of each individual gene as depicted in the control cells and tissues were subtracted from each individual gene measurement performed in the target tissue. Thereafter, the individual, filtered CT values were normalized to the different reference genes. Differences between the CT values of the quantitative PCR reactions of the ARCHEON genes and the reference genes remaining after filtering the primer/probe

performance differences were determined and transformed into „copy numbers per cell". This was done by subtracting the CT values of the target genes from the CT values of the reference genes. The resulting ΔCT values were then transformed in gene copy numbers, with the ΔCT value of the reference gene (ΔCT=0) being defined as „2 copies per cell", by the following formula: 2*(2^(ΔCT*(-1))). All the calculations were done using standard software (Microsoft ExcelTM).

References

Patents cited

U.S. Pat. No. 4,843,155
U.S. Pat. No. 5,262,31
U.S. Pat. No. 4,683,202
U.S. Pat. No. 5,593,839
U.S. Pat. No. 5,578,832
U.S. Pat. No. 5,556,752
U.S. Pat. No. 5,631,734
U.S. Pat. No. 5,599,695
U.S. Pat. No. 4,683,195
U.S. Pat. No. 5,498,531
U.S. Pat. No. 5,714,331
U.S. Pat. No. 5,641,673
U.S. Pat. No. 5,223,409
U.S. Pat. No. 5,976,813
U.S. Pat. No. 5,283,317
U.S. Pat No. 6,203,987
U.S. Pat.No. 6,379,895
WO 97/29212
WO 97/27317
WO 95/22058
WO 99/12826
WO 97/02357
WO 94/13804
WO 94/10300
WO 97/14028
WO 99/52708
EP 0 785 280
EP 0 799 897

Chomczynski, P.
Liang, P., and Pardee, A. B., 1993
Mullis, K. B., 1987
Haseloff et al.,
Lander, E.,
Beutel et al.

EP 0 728 520

EP 0 721 016

EP 0 321 201

GB2188638B

Publications cited

(1) Gusterson et al., Journal of Clinical Oncology 10, 1049-1056, 1992

(2) Achuthan et al., Cancer Genet Cytogenet. 130:166-72, 2001

(3) Tomasetto et al., FEBS Lett. 373: 245-249, 1995

(4) Pragnell et al., FEBS Lett. 291: 253-258, 1991

(5) Nakamichi et al. 1986

(6) Feo et al., Proc. Nat. Acad. Sci. 86: 6691-6695, 1989

(7) Davies et al.,. Proc. Nat. Acad. Sci. 86: 6691-6695, 1989

(8) Lee et al., Molec. Endocr. 9: 243-254, 1995

(9) Drane et al., Oncogene 15: 3013-3024, 1997

(10) Zhu et al., J. Biol. Chem. 272: 25500-25506, 1997

(11) Yuan et al., Proc. Nat. Acad. Sci. 95: 7939-7944, 1998

(12) Zhu et al., Proc. Nat. Acad. Sci. 96: 10848-10853, 1999

(13) Lee et al., Science 268: 836-844, 1995

(14) McCormick et al., Molec. Cell. Biol. 16: 5792-5800, 1996

(15) Tamimi et al., Genomics 40: 355-357, 1997

(16) Valle et al., *FEBS Lett.* 415: 163-168, 1997

(17) Kaneda et al., J. Biol. Chem. 263: 7672-7677, 1988

(18) Hoehe et al., Hum. Molec. Genet. 1: 175-178, 1992

(19) Yang-Feng et al., Abstract Cytogenet. Cell Genet. 40: 784, 1985

(20) Coussens et al., Science 230: 1132-1139, 1985

(21) van de Vijver et al., New Eng. J. Med. 319: 1239-1245, 1988

(22) Slamon et al., Science 244: 707-712, 1989

(23) Fukushige et al., Res. Commun. 134: 477-483, 1986

(24) Kaneko et al., Jpn. J. Cancer Res. 78: 16-19, 1987

(25) Di Fiore et al., Science 237: 178-182, 1987

(26) Popescu et al., Genomics 4: 362-366, 1989

(27) Qiu et al., Nature 393: 83-85, 1998

(28) Yu et al., Molec. Cell 2: 581-591, 1998

(29) Doherty et al., Proc. Nat. Acad. Sci. 96: 10869-10874, 1999

(30) Slamon et al., New Eng. J. Med. 344: 783-792, 2001

(31) Margolis et al., J. Clin. Invest. 102: 821-827, 1998

(32) Tanaka et al., J. Clin. Invest. 102: 821-827, 1998

(33) Dong et al., J. Biol. Chem. 272: 29104-29112, 1997

(34) Stein et al., EMBO J.13:1331-40, 1994

(35) Nagata et al., Nature 319: 415-418, 1986

(36) Le Beau et al., Leukemia 1: 795-799, 1987,

(37) Jansson et al., EMBO J. 2: 561-565, 1983

(38) Thompson et al., Science 237:1610-1614, 1987

(39) Nakai et al., Proc. Nat. Acad. Sci. 85: 2781-2785, 1988

(40) Miyajima et al., Cell 57: 31-39, 1989

(41) Debuire et al., Science 224: 1456-1459, 1984

(42) Petkovich et al., Nature 330: 444-450, 1987

(43) Mattei et al., Hum. Genet. 80: 186-188, 1988.

(44) Williams et al., Molec. Cell. Biol. 18: 2758-2767, 1998

(45) Saha et al, Molec. Cell. Biol. 18: 2758-2767, 1998

(46) Yan et al., Proc. Nat. Acad. Sci. 95: 3603-3608, 1998

(47) Singh et al., Nucleic Acids Res. 16: 3919-3929, 1988

(48) Tsai-Pflugfelder et al., Proc. Nat. Acad. Sci. 85: 7177-7181, 1988

(49) Chung et al., Proc. Nat. Acad. Sci. 86: 9431-9435, 1989

(50) Lang et al., Gene 221: 255-266, 1998

(51) Keith et al., Genes Chromosomes Cancer 4: 169-175, 1992

(52) Kingsmore et al., Mammalian Genome 4: 288-289, 1993

(53) Watt et al., Biochem. J. 303: 681-695, 1994

(54) Kiefer et al., J. Biol. Chem. 267: 12692-12699, 1992

(55) Shimasaki et al., Molec. Endocr. 4: 1451-1458, 1990

(56) Zazzi et al., Genomics 49: 401-410, 1998

(57) Bajalica et al., Hum. Genet. 89: 234-236, 1992

(58) Tonin et al., Genomics 18: 414-417, 1993

(59) Birkenbach et al., J. Virol. 67: 2209-2220, 1993

(60) Wang et al., Proc. Nat. Acad. Sci. 95: 492-498, 1998

(61) Klochendler-Yeivin et al., Curr Opin Genet Dev 121:73-9, 2002

(62) Ring et al., Genomics 51:140-3, 1998

(63) Darmon et al., Molec. Biol. Rep. 12: 277-283, 1987

(64) Zhou et al., J. Biol. Chem. 263: 15584-15589, 1988

(65) Korge et al., Proc. Nat. Acad. Sci. 89: 910-914, 1992

(66) Lessin et al., J. Invest. Derm. 91: 572-578, 1988

(67) Romano et al., Cytogenet. Cell Genet. 58: 2009-2010, 1991.

(68) Fuchs et al., Proc. Nat. Acad. Sci. 89: 6906-6910, 1992

(69) Rogaev et al., Nature Genet. 5: 158-162, 1993

(70) Liu et al., Curr. Eye Res. 12: 963-974, 1993

(71) Nishida et al., Invest. Ophthal. Vis. Sci. 37: 1800-1809, 1996

(72) Nishida et al., Am. J. Hum. Genet. 61: 1268-1275, 1997

(73) Meesmann and Wilke, 1939

(74) Corden et al., Exp. Eye Res. 70: 41-49, 2000

(75)van de Vijver et al., Mol Cell Biol 7, 2091-23, 1987

(76)Offterdinger et al., Biochem Biophys Res Comm 251, 907-13, 1988

(77) Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989

(78) Ausubel et al., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y., 1989.

(79) Tedder, T. F. et al., Proc. Natl. Acad. Sci. U.S.A. 85:208-212, 1988

(80) Hedrick, S. M. et al., Nature 308:149-153, 1984

(81) Lee, S. W. et al., Proc. Natl. Acad. Sci. U.S.A. 88:4225, 1984

(82) Sarkar, PCR Methods Applic. 2, 318-322, 1993

(83) Triglia et al., Nucleic Acids Res. 16, 81-86, 1988

(84) Lagerstrom et al., PCR Methods Applic. 1, 111-119, 1991

(85) Copeland & Jenkins, Trends in Genetics 7: 113-118, 1991

(86) Cohen, et al., Nature 366: 698-701, 1993

(87) Bonner et al., J. Mol. Biol. 81, 123 1973

(88) Bolton and McCarthy, Proc. Natl. Acad. Sci. U.S.A. 48, 1390 1962

(89) Plump et al., Cell 71: 343-353, 1992

(90) Altschul et al., Bull. Math. Bio. 48:603, 1986,

(91) Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915, 1992

(92) Pearson & Lipman, Proc. Nat'l Acad. Sci. USA 85:2444, 1988

(93) Pearson et al., Meth. Enzymol. 183:63, 1990

(94) Needleman & Wunsch, J. Mol. Biol.48:444, 1970

(95) Sellers, SIAM J. Appl. Math.26:787, 1974

(96) Takamatsu, EMBO J. 6, 307-311, 1987

(97) Coruzzi et al., EMBO J. 3, 1671-1680, 1984

(98) Broglie et al., Science 224, 838-843, 1984

(99) Winter et al., Results Probl. Cell Differ. 17, 85-105, 1991

(100) Engelhard et al., Proc. Nat. Acad. Sci. 91, 3224-3227, 1994

(101) Logan & Shenk, Proc. Natl. Acad. Sci. 81, 3655-3659, 1984

(102) Scharf et al., Results Probl. Cell Differ. 20, 125-162, 1994

(103) Freshney R.I., ed., ANIMAL CELL CULTURE, 1986

(104) Wigler et al., Cell 11, 223-232, 1977

(105) Lowy et al., Cell 22, 817-823, 1980

(106) Wigler et al., Proc. Natl. Acad. Sci. 77, 3567-3570, 1980

(107) Colbere-Garapin et al., J. Mol. Biol. 150, 114, 1981

(108) Hartman & Mulligan, Proc. Natl. Acad. Sci. 85, 8047-8051, 1988

(109) Rhodes et al., Methods Mol. Biol. 55, 121-131, 1995

(110) Hampton et al., SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990

(111) Maddox et al., J. Exp. Med. 158, 1211-1216, 1983

(112) Porath et al., Prot. Exp. Purif. 3, 263-281, 1992

(113) Kroll et al., DNA Cell Biol. 12, 441-453, 1993

(114) Caruthers et al., Nucl. Acids Res. Symp. Ser. 215-223, 1980

(115) Horn et al. Nucl. Acids Res. Symp. Ser. 225-232, 1980

(116) Merrifield, J. Am. Chem. Soc. 85, 2149-2154, 1963

(117) Roberge et al., Science 269, 202-204, 1995

(118) Creighton, PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES, WH and Co., New York, N.Y., 1983

(119) Cronin et al., Human Mutation 7:244, 1996

(120) Landegran et al., Science 241:1077-1080, 1988

(121) Nakazawa et al., PNAS 91:360-364, 1994

(122) Abravaya et al., *Nuc Acid Res* 23:675-682, 1995

(123) Guatelli, J.C. *et* al., Proc. Natl. Acad. Sci. USA 87:1874-1878, 1990

(124) Kwoh, D.Y. et al., Proc. Natl. Acad. Sci. USA 86:1173 –1177, 1989

(125) Lizardi, P.M. *et* al., Bio/Technology 6:1197, 1988

(126) Brown, Meth. Mol. Biol. 20, 18, 1994

(127) Sonyeaux, Meth. Mol. Biol. 26, 1-72, 1994

(128) Uhlmann et al., Chem. Rev. 90, 543-583, 1990

(129) Gee et al., in Huber & Carr, MOLECULAR AND IMMUNOLOGIC APPROACHES, Publishing Co., Mt. Kisco, N.Y., 1994

(130) Agrawal et al., Trends Biotechnol. 10, 152-158, 1992

(131) Uhlmann et al., Tetrahedron. Lett. 215, 3539-3542, 1987

(132) Cech, Science 236, 1532-1539, 1987

(133) Cech, Ann. Rev. Biochem. 59, 543-568, 1990

(134) Couture & Stinchcomb, Trends Genet. 12, 510-515, 1996

(135) Haseloff et al. Nature 334, 585-591, 1988

(136) Kohler et al., Nature 256, 495-497, 1985

(137) Kozbor et al., J. Immunol. Methods 81, 3142, 1985

(138) Cote et al., Proc. Natl. Acad. Sci. 80, 2026-2030, 1983

(139) Cole et al., Mol. Cell Biol. 62, 109-120, 1984

(140) Morrison et al., Proc. Natl. Acad. Sci. 81, 6851-6855, 1984

(141) Neuberger et al., Nature 312, 604-608, 1984

(142) Takeda et al., Nature 314, 452-454, 1985

(143) Burton, Proc. Natl. Acad. Sci. 88, 11120-11123, 1991

(144) Thirion et al., Eur. J. Cancer Prev. 5, 507-11, 1996

(145) Coloma & Morrison, Nat. Biotechnol. 15, 159-63, 1997

(146) Mallender & Voss, J. Biol. Chem. 269, 199-206, 1994

(147) Verhaar et al., Int. J. Cancer 61, 497-501, 1995

(148) Nicholls et al., J. Immunol. Meth. 165, 81-91, 1993

(149) Orlandi et al., Proc. Natl. Acad. Sci. 86, 3833-3837, 1989

(150) Winter et al., Nature 349, 293-299, 1991

(151) Lam, Anticancer Drug Des. 12, 145, 1997

(152) DeWitt et al., Proc. Natl. Acad. Sci. U.S.A. 90, 6909, 1993

(153) Erb et al. Proc. Natl. Acad. Sci. U.S.A. 91, 11422, 1994

(154) Zuckermann et al., J. Med. Chem. 37, 2678, 1994

(155) Cho et al., Science 261, 1303, 1993

(156) Carell et al., Angew. Chem. Int. Ed. Engl. 33, 2059 & 2061, 1994

(157) Gallop et al., J. Med. Chem. 37, 1233, 1994

(158) Houghten, BioTechniques 13, 412-421, 1992

(159) Lam, Nature 354, 8284, 1991

(160) Fodor, Nature 364, 555-556, 1993

(161) Cull et al., Proc. Natl. Acad. Sci. U.S.A. 89, 1865-1869, 1992

(162) Scott & Smith, Science 249, 386-390, 1990

(163) Devlin, Science 249, 404-406, 1990

(164) Cwirla et al., Proc. Natl. Acad. Sci. 97, 6378-6382, 1990

(165) Felici, J. Mol. Biol. 222, 301-310, 1991

(166) Jayawickreme et al., Proc. Natl. Acad. Sci. U.S.A. 19, 1614-1618, 1994

(167) Chelsky, Strategies for Screening Combinatorial Libraries 1995

(168) Salmon et al., Molecular Diversity 2, 57-63, 1996

(169) McConnell et al., Science 257, 1906-1912, 1992

(170) Sjolander & Urbaniczky, Anal. Chem. 63, 2338-2345, 1991

(171) Szabo et al., Curr. Opin. Struct. Biol. 5, 699-705, 1995

(172) Zervos et al., Cell 72, 223-232, 1993

(173) Madura et al., J. Biol. Chem. 268, 12046-12054, 1993

(174) Bartel et al., BioTechniques 14, 920-924, 1993

(175) Iwabuchi et al., Oncogene 8, 1693-1696, 1993

(176) Findeis et al. Trends in Biotechnol. 11, 202-205, 1993

(177) Chiou et al., GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE TRANSFER J.A. Wolff, ed., 1994

(178) Wu & Wu, J. Biol. Chem. 263, 621-24, 1988

(179) Wu et al., J. Biol. Chem. 269, 542-46, 1994

(180) Zenke et al., Proc. Natl. Acad. Sci. U.S.A. 87, 3655-59, 1990

(181) Wu et al., J. Biol. Chem. 266, 338-42, 1991

(182) REMINGTON'S PHARMACEUTICAL SCIENCES Maack Publishing Co., Easton, Pa.

(183) Hille, Excitable Membranes, Sunderland, MA, Sinauer Associates, Inc.

Table 1

DNA SEQ ID NO:	Protein SEQ ID NO:	Genbank ID	Unigene_v162_ID	Locus Link ID	Gene Name
1	27	NM_006148.1	Hs.334851	3927	LASP1
2	28	NM_000723.1	Hs.635	782	CACNB1
3	29	NM_000981.1	Hs.381061	6143	RPL19
4	30	Y13467	Hs.15589	5469	PPARGBP
5	31	NM_016507.1	Hs.416108	51755	CrkRS/CRK7
6	32	AB021742.1	Hs:322431	4761	NEUROD2
7	33	NM_006804.1	Hs.77628	10948	MLN64/STARD3
8	34	NM_003673.1	Hs.343603	8557	TELETHONIN
9	35	NM_002686.1	Hs.1892	5409	PNMT
10	36	X03363.1	Hs.446352	2064	ERBB2
11	37	AB008790.1	Hs.86859	2886	GRB7
12	38	NM_002809.1	Hs.9736	5709	PSMD3
13	39	NM_000759.1	Hs.2233	1440	GCSFG/CSF3
14	40	AI023317 NM_014815	Hs.23106	9862	KIAA0130/ TRAP100
15	41	X55005	Hs.724	7067	c-erbA-1 /THRA
16	42	X72631	Hs.2769166	9572	NR1D1
17	43	NM_007359.1	Hs.83422	22794	MLN51
18	44	U77949.1	Hs.405958	990	CDC6
19	45	U41742.1 NM_000964	Hs.361071	5914	RARA
20	46	NM_001067.1	Hs.156346	7153	TOP2A
21	47	NM_001552.1	Hs.1516	3487	IGFBP4
22	48	NM_001838.1	Hs.1652	1236	CCR7 EBI1
23	49	NM_003079.1	Hs.437546	6605	SMARCE1 BAF57
24	50	X14487	Hs.99936	3858	KRT10
25	51	NM_000223.1	Hs.66739	3859	KRT12
26	52	NM_002279.2	Hs.32950	3884	/KRTHA3B
53	76	NM_005937	Hs.497128	4302	MLLT6

Table 1 (continued)

DNA SEQ ID NO:	Protein SEQ ID NO:	Genbank ID	Unigene_v162_ID	Locus Link ID	Gene Name
54	77	XM_008147/ NM_007144	Hs.371617	7703	ZNF144/RNF110
55	78	NM_138687	Hs.9605	8396	PIP5K2B
56	79	NM_020405	Hs.125036	57125	TEM7/PLXDC1
57	80	AF129512	Hs.258579	22806	ZNFN1A3
58	81	XM_085731 NM_133264	Hs.421622	147179	WIRE
59	82	NM_002795	Hs.82793	5691	PSMB3
60	83	NM_033419	Hs.91668	93210	MGC9753 Variant a /CAB2
61	84	NM_033419	Hs.91668	93210	MGC9753 Variant c
62	85	NM_033419	Hs.91668	93210	MGC9753 Variant d
63	86	NM_033419	Hs.91668	93210	MGC9753 Variant e
64	87	NM_033419	Hs.91668	93210	MGC9753 Variant g
65	88	NM_033419	Hs.91668	93210	MGC9753 Variant h
66	89	NM_033419	Hs.91668	93210	MGC9753 Variant i
67	90	AF395708	Hs.133167	94103	ORMDL3
68	91	NM_032875	Hs.194498	84961	MGC15482
69	92	NM_032192	Hs.286192	84152	PPP1R1B
70	93	NM_032339	Hs.333526	84299	MGC14832
71	94	NM_057555 NM_139280	Hs.133167	51242	LOC51242 /ORMDL3
72	95	NM_017748	Hs.406223	54883	FLJ20291
73	96	NM_018530	Hs.306777	55876	Pro2521
74	97	NM_016339	Hs.158530	51195	Link-GEFII

Table 1 (continued)

DNA SEQ ID NO:	Protein SEQ ID NO:	Genbank ID	Unigene_v162_ID	Locus Link ID	Gene Name
75	98	NM_032865	Hs.99037	84951	CTEN
315	393	XM_294897	Hs.270564	30837	NAP4
316	394	NM_032351	Hs.19347	84311	MRLP45
317	395	NM_000458	Hs.408093	6928	TCF2
318	396	NM_152300	Hs.380430	11056	ROK1
319	397	NM_019010	Hs.84905	54474	KRT20
320	398	NM_173213	Hs.9029	25984	KRT23
321	399	NM_033185	Hs.307025	85293	KRTAP3-3
322	400	NM_031959	Hs.307026	83897	KRTAP3-2
323	401	NG_000941		85345	KRTAP3P1
324	402	NM_031958	Hs.307027	83896	KRTAP3-1
325	403	NM_031957	Hs.307030	83895	KRTAP1-5
326	404	NM_030966	Hs.247935	81850	KRTAP1-3
327	405	NM_030967	Hs.247934	81851	KRTAP1-1
328	406	AJ302536		85296	KRTAP2-2
329	407	NM_033184		85294	KRTAP2-4
330	408	NG_000939		85343	KRTAP2P1
331	409	NM_033061	Hs.380164	85287	KRTAP4-7
332	410	NM_033059	Hs.307015	85282	KRTAP4-14
333	411	NM_031854	Hs.307016	83755	KRTAP4-12
334	412	NM_033188	Hs.307016	83755	KRTAP4-5
335	413	NM_033186		85283	KRTAP4-13
336	414	NM_032524	Hs.307022	84616	KRTAP4-4
337	415	NM_033062	Hs.380165	85291	KRTAP4-2
338	416	NM_033060	Hs.380165	85291	KRTAP4-10
339	417	NM_031961	Hs.307013	83899	KRTAP9-2
340	418	NM_031962	Hs.307012	83900	KRTAP9-3
341	419	NM_031963	Hs.307011	83901	KRTAP9-8
342	420	NM_030975	Hs.307010	81870	KRTAP9-9
343	421	NM_033191		85280	KRTAP9-4
344	422	NG_000942		85347	KRTAP9P1
345	423	XM_210345	Hs.463016	85276	KRTAP16-1
346	424	NM_031964	Hs.307009	83902	KRTAP17-1

Table 1 (continued)

DNA SEQ ID NO:	Protein SEQ ID NO:	Genbank ID	Unigene_v162_ID	Locus Link ID	Gene Name
347	425	NM_004138	Hs.197874	3883	KRTHA3A
348	426	NM_002279	Hs.32950	3884	KRTHA3B
349	427	NM_021013	Hs.296942	3885	KRTHA4
350	428	NM_002277	Hs.41696	3881	KRTHA1
351	429	Y16795		8686	KRTHAP1
352	430	NM_003770	Hs.159403	8688	KRTHA7
353	431	NM_006771	Hs.248188	8687	KRTHA8
354	432	NM_002278	Hs.41752	3882	KRTHA2
355	433	NM_002280	Hs.73082	3886	KRTHA5
356	434	NM_003771	Hs.248189	8689	KRTHA6
357	435	NM_002274	Hs.433871	3860	KRT13
358	436	NM_002275	Hs.80342	3866	KRT15
359	437	NM_002276	Hs.309517	3880	KRT19
360	438	NM_000226	Hs.2783	3857	KRT9
361	439	NM_000526	Hs.355214	3861	KRT14
362	440	NM_005557	Hs.432448	3868	KRT16
363	441	NM_000422	Hs.2785	3872	KRT17
364	442	NM_005556	Hs.23881	3855	KRT7
365	443	NG_000944		85349	KRTHBP4
366	444	NG_000943		85348	KRTHBP3
367	445	NM_002281	Hs.170925	3887	KRTHB1
368	446	NM_002284	Hs:278658	3892	KRTHB6
369	447	NM_002282	Hs.182506	3889	KRTHB3
370	448	NG_000940		85344	KRTHBP2
371	449	NM_002283	Hs.182507	3891	KRTHB5
372	450	NM_033045	Hs.272336	3890	KRTHB4
373	451	NM_033033	Hs.134640	3888	KRTHB2
374	452	Y19213		85340	KRTHBP1
375	453	NM_005555	Hs.432677	3854	KRT6B
376	454	NM_173086	Hs.446417	286887	KRT6E
377	455	NM_058242		140446	KRT6C
378	456	NM_005554	Hs.367762	3853	KRT6A
379	457	NM_000424	Hs.433845	3852	KRT5

Table 1 (continued)

DNA SEQ ID NO:	Protein SEQ ID NO:	Genbank ID	Unigene_v162_ID	Locus Link ID	Gene Name
380	458	NM_033448	Hs.55278	112802	KRT6IRS
381	459	NM_175053	Hs.56255	121391	KRT6IRS4
382	460	NM_080747/ AY033495	Hs.147040	140807	K6IRS2/ KRT6
383	461	NM_175068	Hs.319101	55410	KRT6IRS3
384	462	NM_000423	Hs.707	3849	KRT2A
385	463	NM_006121	Hs.80828	3848	KRT1
386	464	NM_057088	Hs.410397	3850	KRT3
387	465	NM_002272	Hs.371139	3851	KRT4
388	466	NM_002273	Hs.356123	3856	KRT8
389	467	NM_000224	Hs.406013	3875	KRT18
390	468	NM_032950	Hs.380710	79148	MMP28
391	469	NM_005419	Hs.72988	6773	STAT2
392	470	NM_002046	Hs.169476	2597	GAPDH

Table 2

DNA SEQ ID NO:	Gene description
1	Member of a subfamily of LIM proteins that contains a LIM domain and an SH3 (Src homology region 3) domain
2	Beta 1 subunit of a voltage-dependent calcium channel (dihydropyridine receptor), involved in coupling of excitation and contraction in muscle, also acts as a calcium channel in various other tissues
3	Ribosomal protein L19, component of the large 60S ribosomal subunit
4	Protein with similarity to nuclear receptor-interacting proteins; binds and co-activates the nuclear receptors PPARalpha (PPARA), RARalpha (RARA), RXR, TRbeta1, and VDR
5	we26e02.x1 CDC2-related protein kinase 7
6	Neurogenic differentiation, a basic-helix-loop-helix transcription factor that mediates neuronal differentiation
7	Protein that is overexpressed in malignant tissues, contains a putative trans-membrane region and a StAR Homology Domain (SHD), may function in steroidogenesis and contribute to tumor progression
8	Telethonin, a sarcomeric protein specifically expressed in skeletal and heart muscle, caps titin (TTN) and is important for structural integrity of the sarcomere
9	Phenylethanolamine N-methyltransferase, acts in catecholamine biosynthesis to convert norepinephrine to epinephrine
10	Tyrosine kinase receptor that has similarity to the EGF receptor, a critical component of IL-6 signaling through the MAP kinase pathway, overexpression associated with prostate, ovary and breast cancer
11	Growth factor receptor-bound protein, an SH2 domain-containing protein that has isoforms which may have a role in cell invasion and metastatic progression of esophageal carcinomas
12	Non-ATPase subunit of the 26S proteasome (prosome, macropain)
13	Granulocyte colony stimulating factor, a glycoprotein that regulates growth, differentiation, and survival of neutrophilic granulocytes

Table 2 (continued)

DNA SEQ ID NO:	Gene description
14	Member of the Vitamin D Receptor Interacting Protein co-activator complex, has strong similarity to thyroid hormone receptor-associated protein (murine Trap100) which function as a transcriptional coregulator
15	Thyroid hormone receptor alpha, a high affinity receptor for thyroid hormone that activates transcription; homologous to avian erythroblastic leukemia virus oncogene
16	encoding Rev-ErbAalp nuclear receptor subfamily 1, group D, member 1
17	Protein that is overexpressed in breast carcinomas
18	Protein which interacts with the DNA replication proteins PCNA and Orc1, translocates from the nucleus following onset of S phase; S. cerevisiae homolog Cdc6p is required for initiation of S phase
19	Retinoic acid receptor alpha, binds retinoic acid and stimulates transcription in a ligand-dependent manner
20	DNA topoisomerase II alpha, member of a family of proteins that relieves torsional stress created by DNA replication, transcription, and cell division;
21	Insulin-like growth factor binding protein, the major IGFBP of osteoblast-like cells, binds IGF1 and IGF2 and inhibits their effects on promoting DNA and glycogen synthesis in osteoblastic cells
22	HUMEBI103 G protein-coupled receptor (EBI 1) gene exon 3 chemokine (C-C motif) receptor 7 G protein-coupled receptor
23	Protein with an HMG 1/2 DNA-binding domain that is subunit of the SNF/SWI complex associated with the nuclear matrix and implicated in regulation of transcription by affecting chromatin structure
24	Keratin 10, a type I keratin that is a component of intermediate filaments and is expressed in terminally differentiated epidermal cells; mutation of the corresponding gene causes epidermolytic hyperkeratosis
25	Keratin 12, a component of intermediate filaments in corneal epithelial cells; mutation of the corresponding gene causes Meesmann corneal dystrophy
26	Hair keratin 3B, a type I keratin that is a member of a family of structural proteins that form intermediate filaments
53	MLLT6 Myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog,

Table 2 (continued)

DNA SEQ ID NO:	Gene description
	Drosophila); translocated to, 6
54	zinc finger protein 144 (Mel-18)
55	Phosphatidylinositol-4-phosphate 5-kinase type II beta isoform a
56	tumor endothelial marker 7 precursor
57	zinc finger protein, subfamily 1A, 3
58	WASP-binding protein putative cr16 and wip like protein similar to Wiskott-Aldrich syndrome protein
59	Proteasome (prosome, macropain) subunit, beta type, 3
60	Predicted
67	ORM1-like 3 (S. cerevisiae)
68	F-box domain A Receptor for Ubiquitination Targets
69	protein phosphatase 1, regulatory (inhibitor) subunit 1B (dopamine and cAMP regulated phosphoprotein, DARPP-32)
70	Predicted Protein
71	Predicted Protein
72	Predicted Protein
73	Predicted Protein
74	Link-GEFII: Link guanine nucleotide exchange factor II
75	C-terminal tensin-like
315	Homo sapiens Nck, Ash and phospholipase C binding protein (NAP4)
316	Homo sapiens mitochondrial ribosomal protein L45 (MRPL45), nuclear gene encoding mitochondrial protein
317	Homo sapiens transcription factor 2, hepatic; LF-B3; variant hepatic nuclear factor (TCF2), transcript variant a
318	Homo sapiens DEAD (Asp-Glu-Ala-Asp) box polypeptide 52 (DDX52)
319	Homo sapiens keratin 20 (KRT20), ,is a component of intermediate filament network
320	Homo sapiens keratin 23 (histone deacetylase inducible) (KRT23), is a component of intermediate filament network transcript variant 2
321	Homo sapiens keratin associated protein 3-3 (KRTAP3-3)), is a component of intermediate filament network
322	Homo sapiens keratin associated protein 3-2 (KRTAP3-2) ,is a component of

Table 2 (continued)

DNA SEQ ID NO:	Gene description
	intermediate filament network
323	Homo sapiens keratin associated protein 3 pseudogene 1 (KRTAP3P1) on chromosome 17 ,is a component of intermediate filament network
324	Homo sapiens keratin associated protein 3-1 (KRTAP3-1) ,is a component of intermediate filament network
325	Homo sapiens keratin associated protein 1-5 (KRTAP1-5) ,is a component of intermediate filament network
326	Homo sapiens keratin associated protein 1-3 (KRTAP1-3) ,is a component of intermediate filament network
327	Homo sapiens keratin associated protein 1-1 (KRTAP1-1) ,is a component of intermediate filament network
328	HSA302536 Homo sapiens partial mRNA for keratin associated protein KAP2.2 (KRTAP2.2 gene) ,is a component of intermediate filament network
329	Homo sapiens keratin associated protein 2-4 (KRTAP2-4) ,is a component of intermediate filament network
330	Homo sapiens keratin associated protein 2 pseudogene 1 (KRTAP2P1) on chromosome 17,is a component of intermediate filament network
331	Homo sapiens keratin associated protein 4-7 (KRTAP4-7) ,is a component of intermediate filament network
332	Homo sapiens keratin associated protein 4-14 (KRTAP4-14) ,is a component of intermediate filament network
333	Homo sapiens keratin associated protein 4-12 (KRTAP4-12) ,is a component of intermediate filament network
334	Homo sapiens keratin associated protein 4-5 (KRTAP4-5) ,is a component of intermediate filament network
335	Homo sapiens keratin associated protein 4-13 (KRTAP4-13) ,is a component of intermediate filament network
336	Homo sapiens keratin associated protein 4-4 (KRTAP4-4) ,is a component of intermediate filament network
337	Homo sapiens keratin associated protein 4-2 (KRTAP4-2) ,is a component of intermediate filament network
338	Homo sapiens keratin associated protein 4-10 (KRTAP4-10) ,is a component

Table 2 (continued)

DNA SEQ ID NO:	Gene description
	of intermediate filament network
339	Homo sapiens keratin associated protein 9-2 (KRTAP9-2) ,is a component of intermediate filament network
340	Homo sapiens keratin associated protein 9-3 (KRTAP9-3) ,is a component of intermediate filament network
341	Homo sapiens keratin associated protein 9-8 (KRTAP9-8) ,is a component of intermediate filament network
342	Homo sapiens keratin associated protein 9-9 (KRTAP9-9) ,is a component of intermediate filament network
343	Homo sapiens keratin associated protein 9-4 (KRTAP9-4) ,is a component of intermediate filament network
344	Homo sapiens keratin associated protein 9 pseudogene 1 (KRTAP9P1) on chromosome 17 ,is a component of intermediate filament network
345	Homo sapiens keratin associated protein 16-1 (KRTAP16-1) ,is a component of intermediate filament network
346	Homo sapiens keratin associated protein 17-1 (KRTAP17-1) ,is a component of intermediate filament network
347	Homo sapiens keratin, hair, acidic, 3A (KRTHA3A) ,is a component of intermediate filament network
348	Homo sapiens keratin, hair, acidic, 3B (KRTHA3B) ,is a component of intermediate filament network
349	Homo sapiens keratin, hair, acidic, 4 (KRTHA4) ,is a component of intermediate filament network
350	Homo sapiens keratin, hair, acidic, 1 (KRTHA1) ,is a component of intermediate filament network
351	HSA16795 Homo sapiens KRTHAP1 pseudogene,is a component of intermediate filament network
352	Homo sapiens keratin, hair, acidic, 7 (KRTHA7) ,is a component of intermediate filament network
353	Homo sapiens keratin, hair, acidic, 8 (KRTHA8) ,is a component of intermediate filament network
354	Homo sapiens keratin, hair, acidic, 2 (KRTHA2) ,is a component of

Table 2 (continued)

DNA SEQ ID NO:	Gene description
	intermediate filament network
355	Homo sapiens keratin, hair, acidic, 5 (KRTHA5) ,is a component of intermediate filament network
356	Homo sapiens keratin, hair, acidic, 6 (KRTHA6) ,is a component of intermediate filament network
357	Homo sapiens keratin 13 (KRT13), transcript variant 2,is a component of intermediate filament network
358	Homo sapiens keratin 15 (KRT15) ,is a component of intermediate filament network
359	Homo sapiens keratin 19 (KRT19) ,is a component of intermediate filament network
360	Homo sapiens keratin 9 (epidermolytic palmoplantar keratoderma) (KRT9) ,is a component of intermediate filament network
361	Homo sapiens keratin 14 (epidermolysis bullosa simplex, Dowling-Meara, Koebner) (KRT14) ,is a component of intermediate filament network
362	Homo sapiens keratin 16 (focal non-epidermolytic palmoplantar keratoderma) (KRT16) ,is a component of intermediate filament network
363	Homo sapiens keratin 17 (KRT17) ,is a component of intermediate filament network
364	Homo sapiens keratin 7 (KRT7) ,is a component of intermediate filament network
365	Homo sapiens psihHbD hair keratin pseudogene (KRTHBP4) on chromosome 12,is a component of intermediate filament network
366	Homo sapiens psihHbC hair keratin pseudogene (KRTHBP3) on chromosome 12,is a component of intermediate filament network
367	Homo sapiens keratin, hair, basic, 1 (KRTHB1) ,is a component of intermediate filament network
368	Homo sapiens keratin, hair, basic, 6 (monilethrix) (KRTHB6) ,is a component of intermediate filament network
369	Homo sapiens keratin, hair, basic, 3 (KRTHB3) ,is a component of intermediate filament network
370	Homo sapiens psihHbB hair keratin pseudogene (KRTHBP2) on chromosome

Table 2 (continued)

DNA SEQ ID NO:	Gene description
	12 ,is a component of intermediate filament network
371	Homo sapiens keratin, hair, basic, 5 (KRTHB5) ,is a component of intermediate filament network
372	Homo sapiens keratin, hair, basic, 4 (KRTHB4), ,is a component of intermediate filament network
373	Homo sapiens keratin, hair, basic, 2 (KRTHB2) ,is a component of intermediate filament network
374	HSPSIHHBA Homo sapiens putative psihHbA pseudogene for hair keratin, exons 2 to 7
375	Homo sapiens keratin 6B (KRT6B) ,is a component of intermediate filament network
376	Homo sapiens keratin 6E (KRT6E),is a component of intermediate filament network
377	Homo sapiens keratin 6C (KRT6C) ,is a component of intermediate filament network
378	Homo sapiens keratin 6A (KRT6A), ,is a component of intermediate filament network
379	Homo sapiens keratin 5 (epidermolysis bullosa simplex, Dowling-Meara/Kobner/Weber-Cockayne types) (KRT5) ,is a component of intermediate filament network
380	Homo sapiens keratin 6 irs (KRT6IRS) ,is a component of intermediate filament network
381	Homo sapiens keratin 6 irs4 (K6IRS4) ,is a component of intermediate filament network
382	Homo sapiens keratin protein K6irs (K6IRS2) ,is a component of intermediate filament network
383	Homo sapiens keratin protein K6irs (K6IRS2) ,is a component of intermediate filament network
384	Homo sapiens keratin 2A (epidermal ichthyosis bullosa of Siemens) (KRT2A) ,is a component of intermediate filament network
385	Homo sapiens keratin 1 (epidermolytic hyperkeratosis) (KRT1) ,is a component of intermediate filament network

Table 2 (continued)

DNA SEQ ID NO:	Gene description
386	Homo sapiens keratin 3 (KRT3) ,is a component of intermediate filament network
387	Homo sapiens keratin 4 (KRT4) ,is a component of intermediate filament network
388	Homo sapiens keratin 8 (KRT8) ,is a component of intermediate filament network
389	Homo sapiens keratin 18 (KRT18) ,is a component of intermediate filament network
390	Homo sapiens matrix metalloproteinase 28 (MMP28), transcript variant 2
391	Homo sapiens signal transducer and activator of transcription 2, 113kDa (STAT2)
392	Homo sapiens glyceraldehyde-3-phosphate dehydrogenase (GAPD)

Table 3

DNA SEQ ID NO:	Gene function	Subcellular localization
1	SH3/SH2 adapter protein	-
2	voltage-gated calcium channel membrane fraction Channel [passive transporter]	Plasma membrane
3	RNA binding structural protein of ribosome protein biosynthesis	Cytoplasm
4	transcription co-activator nucleus Pol II transcription	Nucleus
5	-	-
6	transcription factor transcription regulation from Pol II promoter neurogenesis	-
7	mitochondrial transport steroid and lipid metabolism	Cytoplasm
8	structural protein of muscle sarcomere alignment	Cytoplasm
9	phenylethanolamine N-methyltransferase Transferase	-
10	Neu/ErbB-2 receptor receptor signaling protein tyrosine kinase	Plasma membrane
11	SH3/SH2 adapter protein EGF receptor signaling pathway	Cytoplasm
12	26S proteasome Protein degradation Proteasome subunit	Cytoplasm
13	developmental processes positive control of cell proliferation	Extracellular space
14	fatty acid omega-hydroxylase fatty acid omega-hydroxylase	-
15	DNA-binding protein Transcription factor	Nucleus
16	steroid hormone receptor transcription co-repressor	Nucleus
17	-	-
18	nucleotide binding cell cycle regulator DNA replication checkpoint regulation of CDK activity	nucleus

DNA SEQ ID NO:	Gene function	Subcellular localization
19	retinoic acid receptor transcription co-activator transcription factor	nucleus
20	DNA binding DNA topoisomerase (ATP-hydrolyzing)	nucleus
21	skeletal development DNA metabolism signal transduction cell proliferation	
22		plasma membrane
23	chromatin binding transcription co-activator nucleosome disassembly transcription	nucleus nuclear chromosome
24	Cell structure Cytoskeletal Epidermal Development and Maintenance	cytoplasm
25	structural protein vision cell shape and cell size control intermediate filament	cytoplasm
26	cell shape and cell size control Cell structure	cytoplasm
53		-
54	leucine-zipper containing fusion	-
55		
56	Tumor endothelial marker 7 precursor; may be involved in angiogenesis	-
57	Aiolos; DNA binding protein that may be a transcription factor; has strong similarity to murine Znfn1a3, contains zinc finger domain	-
58	The WASP-binding protein WIRE has a role in the regulation of the actin filament system downstream of the platelet-derived growth factor receptor	-
59		-
60		-
61		-

DNA SEQ ID NO:	Gene function	Subcellular localization
67		-
68		-
69	Midbrain dopaminergic neurons play a critical role in multiple brain functions, and abnormal signaling through dopaminergic pathways has been implicated in several major neurologic and psychiatric disorders. One well-studied target for the actions of dopamine is DARPP32.	-
70		-
71		-
72		-
73		-
74	Brain-specific guanine nucleotide exchange factor; activates the ERK/MAP kinase cascade plus R-Ras and H-ras; activates targets through a Ca2+- and diacylglycerol-sensitive mechanism; active protein associates with membranes	-
75	C-terminal tensin-like Phosphotyrosine-binding domain, phosphotyrosine-interaction (PI) domain	-
315		
316		
317		
318		cytoplasm
319	KRT20, integral part of the intermediate filamentous network	Cytoplasm
320	KRT23, integral part of the intermediate filamentous network	Cytoplasm

DNA SEQ ID NO:	Gene function	Subcellular localization
321	KRTAP3-3, integral part of the intermediate filamentous network	Cytoplasm
322	KRTAP3-2, integral part of the intermediate filamentous network	Cytoplasm
323	KRTAP3P1, integral part of the intermediate filamentous network	Cytoplasm
324	KRTAP3-1, integral part of the intermediate filamentous network	Cytoplasm
325	KRTAP1-5, integral part of the intermediate filamentous network	cytoplasm
326	KRTAP1-3, integral part of the intermediate filamentous network	cytoplasm
327	KRTAP1-1, integral part of the intermediate filamentous network	Cytoplasm
328	KRTAP2-2, integral part of the intermediate filamentous network	Cytoplasm
329	KRTAP2-4, integral part of the intermediate filamentous network	Cytoplasm
330	KRTAP2P1, integral part of the intermediate filamentous network	Cytoplasm
331	KRTAP4-7, integral part of the intermediate filamentous network	Cytoplasm
332	KRTAP4-14, integral part of the intermediate filamentous network	Cytoplasm
333	KRTAP4-12, integral part of the intermediate filamentous network	cytoplasm
334	KRTAP4-5, integral part of the intermediate filamentous network	cytoplasm
335	KRTAP4-13, integral part of the intermediate filamentous network	Cytoplasm
336	KRTAP4-4, integral part of the intermediate filamentous network	Cytoplasm
337	KRTAP4-2, integral part of the intermediate filamentous network	Cytoplasm
338	KRTAP4-10, integral part of the intermediate filamentous network	Cytoplasm
339	KRTAP9-2, integral part of the intermediate filamentous network	Cytoplasm
340	KRTAP9-3, integral part of the intermediate filamentous network	Cytoplasm

DNA SEQ ID NO:	Gene function	Subcellular localization
341	KRTAP9-8, integral part of the intermediate filamentous network	cytoplasm
342	KRTAP9-9, integral part of the intermediate filamentous network	cytoplasm
343	KRTAP9-4, integral part of the intermediate filamentous network	Cytoplasm
344	KRTAP9P1, integral part of the intermediate filamentous network	Cytoplasm
345	KRTAP16-1, integral part of the intermediate filamentous network	Cytoplasm
346	KRTAP17-1, integral part of the intermediate filamentous network	Cytoplasm
347	KRTHA3A, integral part of the intermediate filamentous network	Cytoplasm
348	KRTHA3B, integral part of the intermediate filamentous network	Cytoplasm
349	KRTHA4, integral part of the intermediate filamentous network	cytoplasm
350	KRTHA1, integral part of the intermediate filamentous network	cytoplasm
351	KRTHAP1, integral part of the intermediate filamentous network	Cytoplasm
352	KRTHA7, integral part of the intermediate filamentous network	Cytoplasm
353	KRTHA8, integral part of the intermediate filamentous network	Cytoplasm
354	KRTHA2, integral part of the intermediate filamentous network	Cytoplasm
355	KRTHA5, integral part of the intermediate filamentous network	Cytoplasm
356	KRTHA6, integral part of the intermediate filamentous network	Cytoplasm
357	KRT13, integral part of the intermediate filamentous network	cytoplasm
358	KRT15, integral part of the intermediate filamentous network	cytoplasm
359	KRT19, integral part of the intermediate filamentous network	Cytoplasm
360	KRT9, integral part of the intermediate filamentous network	Cytoplasm

DNA SEQ ID NO:	Gene function	Subcellular localization
361	KRT14, integral part of the intermediate filamentous network	Cytoplasm
362	KRT16, integral part of the intermediate filamentous network	Cytoplasm
363	KRT17, integral part of the intermediate filamentous network	Cytoplasm
364	KRT7, integral part of the intermediate filamentous network	Cytoplasm
365	KRTHBP4, integral part of the intermediate filamentous network	cytoplasm
366	KRTHBP3, integral part of the intermediate filamentous network	cytoplasm
367	KRTHB1, integral part of the intermediate filamentous network	Cytoplasm
368	KRTHB6, integral part of the intermediate filamentous network	Cytoplasm
369	KRTHB3, integral part of the intermediate filamentous network	Cytoplasm
370	KRTHBP2, integral part of the intermediate filamentous network	Cytoplasm
371	KRTHB5, integral part of the intermediate filamentous network	Cytoplasm
372	KRTHB4, integral part of the intermediate filamentous network	Cytoplasm
373	KRTHB2, integral part of the intermediate filamentous network	cytoplasm
374	KRTHBP1, integral part of the intermediate filamentous network	cytoplasm
375	KRT6B, integral part of the intermediate filamentous network	Cytoplasm
376	KRT6E, integral part of the intermediate filamentous network	Cytoplasm
377	KRT6C, integral part of the intermediate filamentous network	Cytoplasm
378	KRT6A, integral part of the intermediate filamentous network	Cytoplasm
379	KRT5, integral part of the intermediate filamentous network	Cytoplasm
380	KRT6IRS, integral part of the intermediate filamentous network	Cytoplasm

DNA SEQ ID NO:	Gene function	Subcellular localization
381	KRT6IRS4, integral part of the intermediate filamentous network	cytoplasm
382	KRT6, integral part of the intermediate filamentous network	Cytoplasm
383	KRT6IRS3, integral part of the intermediate filamentous network	Cytoplasm
384	KRT2A, integral part of the intermediate filamentous network	Cytoplasm
385	KRT1, integral part of the intermediate filamentous network	Cytoplasm
386	KRT3, integral part of the intermediate filamentous network	Cytoplasm
387	KRT4, integral part of the intermediate filamentous network	Cytoplasm
388	KRT8, integral part of the intermediate filamentous network	cytoplasm
389	KRT18, integral part of the intermediate filamentous network	Cytoplasm

Table 4

DNA SEQ ID NO:	Protein SEQ ID NO:	Gene Name	DBSNP ID	Type	Codon	AA-Seq
9	34	ERBB2	rs2230698	coding-synon	TCA|TCG	S|S
9	34	ERBB2	rs2230700	noncoding		
9	34	ERBB2	rs1058808	coding-nonsynon	CCC|GCC	P|A
9	34	ERBB2	rs1801200	noncoding		
9	34	ERBB2	rs903506	noncoding		
9	34	ERBB2	rs2313170	noncoding		
9	34	ERBB2	rs1136201	coding-nonsynon	ATC|GTC	I|V
9	34	ERBB2	rs2934968	noncoding		
9	34	ERBB2	rs2172826	noncoding		
9	34	ERBB2	rs1810132	coding-nonsynon	ATC|GTC	I|V
9	34	ERBB2	rs1801201	noncoding		
14	39	c-erbA-1	rs2230702	coding-synon	TCC|TCT	S|S
14	39	c-erbA-1	rs2230701	coding-synon	GCC|GCT	A|A
14	39	c-erbA-1	rs1126503	coding-nonsynon	ACC|AGC	T|S
14	39	c-erbA-1	rs3471	noncoding		
19	44	TOP2A	rs13695	noncoding		
19	44	TOP2A	rs471692	noncoding		
19	44	TOP2A	rs558068	noncoding		
19	44	TOP2A	rs1064288	noncoding		
19	44	TOP2A	rs1061692	coding-synon	GGA|GGG	G|G
19	44	TOP2A	rs520630	noncoding		
19	44	TOP2A	rs782774	coding-nonsynon	AAT|ATT|ATT|TTT	N|I|I|F
19	44	TOP2A	rs565121	noncoding		
19	44	TOP2A	rs2586112	noncoding		
19	44	TOP2A	rs532299	coding-nonsynon	TTT|GTT	F|V

Table 4 (continued)

DNA SEQ ID NO:	Protein SEQ ID NO:	Gene Name	DBSNP ID	Type	Codon	AA-Seq
19	44	TOP2A	rs2732786	noncoding		
19	44	TOP2A	rs1804539	noncoding		
19	44	TOP2A	rs1804538	noncoding		
19	44	TOP2A	rs1804537	noncoding		
19	44	TOP2A	rs1141364	coding-synon	AAA|AAG	K|K
23	48	KRT10	rs12231	noncoding		
23	48	KRT10	rs1132259	coding-nonsynon	CAT|CGT	H|R
23	48	KRT10	rs1132257	coding-synon	CTG|TTG	L|L
23	48	KRT10	rs1132256	coding-synon	GCC|GCT	A|A
23	48	KRT10	rs1132255	coding-synon	CTG|TTG	L|L
23	48	KRT10	rs1132254	coding-synon	GGC|GGT	G|G
23	48	KRT10	rs1132252	coding-synon	TTC|TTT	F|F
23	48	KRT10	rs1132268	coding-nonsynon	CAG|GAG	Q|E
23	48	KRT10	rs1132258	coding-nonsynon	CGG|TGG	R|W

Table 5

PRIMER		SEQUENCE	
CACNB1	FAM 5′	CCATATATAAAACCACTGTCCTGTCCTTTGTGGCT	3′TAMRA
CACNB1FOR	5′	CCCCCATCTGTCTGTCTATATTTGTC	3′
CACNB1REV	5′	TGCCTACGCTGACGACTATGTG	3′
CDC6	FAM 5′	TTTGGTTTTCTACAACTGTTGCTAT	3′TAMRA
CDC6 FOR	5′	GGGCTCCACACACCAGATG	3′
CDC6 REV	5′	ACGCTCTGAGCACCCTCTACA	3′
EBI1-1	FAM 5′	TGTCACAGGGACTGAAAACCTCTCCTCATGT	3′TAMRA
EBI1-1 FOR	5′	CCCAAGGCCACGAGCTT	3′
EBI1-1 REV	5′	TGTTGCTCTCTTAACGAATCGAAA	3′
EBI1-2	FAM 5′	CTGGTCAAACAAACTCTCTGAACCCCTCC	3′TAMRA
EBI1-2 FOR	5′	TGGTGAGGAAAAGCGGACAT	3′
EBI1-2 REV	5′	CTGGCTTGGAGGACAGTGAAG	3′
GCSF	FAM 5′	CCAAGCCCTCCCCATCCCATGTAT	3′TAMRA
GCSF FOR	5′	GAGGTGTCGTACCGCGTTCTA	3′
GCSF REV	5′	CCGTTCTGCTCTTCCCTGTCT	3′
GRB7	FAM 5′	CCAGACCCGCTTCACTGACCTGC	3′TAMRA
GRB7 FOR	5′	CGCCTGTACTTCAGCATGGA	3′
GRB7 REV	5′	GCGGTTCAGCTGGTGGAA	3′
HKA3	FAM 5′	ACCCCGAGGCATCACCACAAATCAT	3′TAMRA
HKA3 FOR	5′	AGTTCTGCCTCTCTGACAACCAT	3′
HKA3 REV	5′	TAGGCTCAGAGTCAGACCCAAAC	3′
MLN50	FAM 5′	CCCTCGTGGGCTTGTGCTCGG	3′TAMRA
MLN50 FOR	5′	AAGCCGCCAGTTCATCTTTTT	3′
MLN50 REV	5′	CTTGTGGTTCAAGTCAAATGTTCAG	3′
MLN64-1	FAM 5′	TCTGCCTGCGCTCTCGTCGGT	3′TAMRA
MLN64-1 FOR	5′	GGGCTGGGCACCTGACTT	3′

Table 5 (continued)

PRIMER		SEQUENCE	
MLN64-1REV	5′	CCCAACAAGGGTCCCAGACT	3′
MLN64-2	FAM 5′	CGGCGCATTGAGCGGCG	3′TAMRA
MLN64-2 FOR	5′	CCCAAGGGACTTCGTGAATG	3′
MLN64-2REV	5′	GGCGATCCCTGATGACAAGTA	3′
PPARBP	FAM 5′	AGCACCAACTGTGAACCAGGTACAATGGC	3′TAMRA
PPARBP FOR	5′	GAGGGAGGCTCTGCTTTGG	3′
PPARBP REV	5′	TCACAACTAGCGGGTGAGGAG	3′
PSMD3	FAM 5′	TGCAGAGGAACGGCGTGAGCG	3′TAMRA
PSMD3 FOR	5′	TGAGGTTTCCTCCCAAATCGTA	3′
PSMD3 REV	5′	CAGCTCAAGGGAAGCTGTCATC	3′
RAR	FAM 5′	CCCCCACATGTTCCCCAAGATGCT	3′TAMRA
RAR FOR	5′	GGAGGCGCTAAAGGTCTACGT	3′
RAR REV	5′	TGATGCTTCGCAGGTCAGTAA	3′
RPL23A	FAM 5′	CTCCTGCCCCTCCTAAAGCTGAAGCC	3′TAMRA
RPL23A FOR	5′	GGACGCGTGGGCTTTTC	3′
RPL23A REV	5′	TGTGGCTGTGGACACCTTTC	3′
RPL19	FAM 5′	CCACAAGCTGAAGGCAGACAAGGCC	3′TAMRA
RPL19 FOR	5′	GCGGATTCTCATGGAACACA	3′
RPL19 REV	5′	GGTCAGCCAGGAGCTTCTTG	3′
NEUROD2	FAM 5′	ACCACCTTGCGCAGGTTGTCCAG	3′TAMRA
NEUROD2 FOR	5′	CGCATGCACGACCTGAAC	3′
NEUROD2 REV	5′	GTCTCGATCTTGGACAGCTTCTG	3′
TELE TELETHONIN	FAM 5′	ACACTGTCCACACGGCCCGAGG	3′TAMRA
TELE TELETHONIN FOR	5′	CTGGGCAGAATGGAAGGATCT	3′
TELE TELETHONIN REV	5′	GGGACTCTAGCAGACCCACACT	3′
PENT PNMT	FAM 5′	CACCCACCTGGATTCCCTGTTC	3′TAMRA
PENT PNMT FOR	5′	CCTTCAGACAGGCGTAGATGATG	3′
PENT PNMT REV	5′	GGGTATTATTTCTTTATTAGGTGCCACTT	3′
HER2/NEU;ERBB2	FAM 5′	TTCCCTAAGGCTTTCAGTACCCAGGATCTG	3′TAMRA

Table 5 (continued)

PRIMER		SEQUENCE	
HER2/NEU;ERBB FOR	5′	CCAGCTTGGCCCTTTCCT	3′
HER2/NEU;ERBB REV	5′	GAATGGGTCGCTTTTGTTCTTAG	3′
KIA0130	FAM 5′	TCACGGACCTCAGCCTGCCCCT	3′TAMRA
KIA0130 FOR	5′	TGGTGAAGGTGTCAGCCATGT	3′
KIA0130 REV	5′	TCAGAGTGCAGCAATGGCTTT	3′
THRA	FAM 5′	ACCTCCTTCCCCAGCTCCCC	3′TAMRA
THRA FOR	5′	GGCAACATCTTACTTGTCCTTTGA	3′
THRA REV	5′	CCAAGGAAGCACAGACAACTATTTC	3′
MLN51	FAM 5′	TCCTCCCTATCCATGGCACTAAACCACTTC	3′TAMRA
MLN51 FOR	5′	TGGGCAAGGGCTCCTATCT	3′
MLN51 REV	5′	GTTACCCCTGGCAGACGTATG	3′
TOP2A	FAM 5′	TGCCTCTGAGTCTGAATCTCCCAAAGAGAGA	3′TAMRA
TOP2A FOR	5′	GAGTAGTTATGTGATTATTTCAGCTCTTGAC	3′
TOP2A REV	5′	TCAAATGTTGTCCCCGAGTCT	3′
KRT10	FAM 5′	CAGAAATTCGGAAGACAGAACTATTGTCATGCC T	3′TAMRA
KRT10 FOR	5′	GATTAGTAACCCATAGCAGTTGAAGGT	3′
KRT10 REV	5′	ATTTACTGACGGTGGTCTGAACATAC	3′
K12 KRT12	FAM 5′	TGACAGACTCCAAATCACAAGCACAGTCAAC	3′TAMRA
K12 KRT12 FOR	5′	TGATGGTTTGGAGGAAAGTTTATTT	3′
K12 KRT12 REV	5′	TTTGGTTGGGTCTTTAGAGGAATC	3′
NR1D1	FAM 5′	TGCCAACCATGCATCAGGTAGCCC	3′TAMRA
NR1D1 FOR	5′	CAGCTCACCTGGCAACTTCA	3′
NR1D1 REV	5′	CCTGATTTTCCCAGCGATGT	3′
HSERBT1	FAM 5′	CGCCGCTCCCGGTTCTGCT	3′TAMRA
HSERBT FOR	5′	TGGCCAAGCGTAAGCTGATT	3′
HSERBT REV	5′	GCTGCAGTGATCGGATCATCT	3′
MLLT6	FAM 5′	CACCATGGAGCCCATCGTGCTG	3′TAMRA
MLLT6 FOR	5′	ATCCCCGAGGTGCAATTTG	3′
MLLT6 REV	5′	AGCGATCATGAGGCACGTACT	3′

Table 5 (continued)

PRIMER		SEQUENCE	
ZNF144	FAM 5′	CCTGCCAGAGATAGGAGACCCAGACAGCT	3′TAMRA
ZNF144 FOR	5′	ATCCCCCTGAGCCTTTTCA	3′
ZNF144 REV	5′	CAGCCTCTGGTCCCACCAT	3′
PIP5K2B	FAM 5′	TGATCATCAATTCCAAACCTCTCCCGAA	3′TAMRA
PIP5K2B FOR	5′	CCCCATGGTGTTCCGAAAC	3′
PIP5K2B REV	5′	TGCCAGGAGCCTCCATACC	3′
TEM7	FAM 5′	CAGCCTTCTAAAACACAATGTATTCATGT	3′TAMRA
TEM7 FOR	5′	CCTGAACTTAATGGTAGAATTCAAAGATC	3′
TEM7 REV	5′	TATTAACACTGAGAATCCATGCAGAGA	3′
ZNFN1A3	FAM 5′	TATCTGGTCTCAGGGATTGCTCCTATGTATTCAG C	3′TAMRA
ZNFN1A3 FOR	5′	CACAGAGCCCTGCTGAAGTG	3′
ZNFN1A3 REV	5′	GCGAGGTCATTGGTTTTTAGAAA	3′
WIRE	FAM 5′	CTGTGATCCGAAATGGTGCCAG	3′TAMRA
WIRE FOR	5′	CCGTCTCCACATCCAAACCT	3′
WIRE REV	5′	ACCCATGCATTCGGTATGGT	3′
PSMB3	FAM 5′	AGTGGCACCTGCGCCGAACAA	3′TAMRA
PSMB3 FOR	5′	CCCCATGGTGACTGATGACTT	3′
PSMB3 REV	5′	CCAGAGGGACTCACACATTCC	3′
MGC9753	FAM 5′	CCAGAAACTTTCCATCCCAAAGGCAGTCT	3′TAMRA
MGC9753 FOR	5′	CTGCCCCACAGGAATAGAATG	3′
MGC9753 REV	5′	AAAAATCCAGTCTGCTTCAACCA	3′
ORMDL3	FAM 5′	AGCTGCCCCAGCTCCACGGA	3′TAMRA
ORMDL3 FOR	5′	TCCCTGATGAGCGTGCTTATC	3′
ORMDL3 REV	5′	TCTCAGTACTTATTGATTCCAAAAATCC	3′
MGC15482	FAM 5′	TCCAGTGGAAGCAACCCCAGTGTTC	3′TAMRA
MGC15482 FOR	5′	CACTTCTAGAGCTACCGTGGAGTCT	3′
MGC15482 REV	5′	CCCTCACTTTGTAACCCTTGCT	3′
PPP1R1B	FAM 5′	CAGCGTGGCGCAACAACCCA	3′TAMRA
PPP1R1B FOR	5′	GGGATTGTTTCGCCACACATA	3′

Table 5 (continued)

PRIMER		SEQUENCE	
PPP1R1B REV	5′	CCGATGTTAAGGCCCATAGC	3′
MGC14832	FAM 5′	TAAAATGTCCGGCCAACATGAGTTCCC	3′TAMRA
MGC14832 FOR	5′	CGCAGTGCCTGGCACAT	3′
MGC14832 REV	5′	GACACCCCCTGACCTATGGA	3′
LOC51242	FAM 5′	CAGTGACCTCTCCCGTTCCCTTGGA	3′TAMRA
LOC51242 FOR	5′	TGGGTCCCTGTGTCCTCTTC	3′
LOC51242 REV	5′	AGGGTCAGGAGGGAGAAAAC	3′
FLJ20291	FAM 5′	CCAGTGCCCACCCGTTAAAGAGTCAA	3′TAMRA
FLJ20291 FOR	5′	TTGTGGGACACTCAGTAACTTTGG	3′
FLJ20291 REV	5′	ACAAGCACTCCCACCGAGAT	3′
PRO2521	FAM 5′	AGTCTGTCCTCACTGCCATCGCCA	3′TAMRA
PRO2521 FOR	5′	AAGCCTCTGGGTTTTCCCTTT	3′
PRO2521 REV	5′	CCCACTGGTGACAGGATGGT	3′
Link-GEFII	FAM 5′	CATCTGACATCTTTCCCGTGGAG	3′TAMRA
Link-GEFII FOR	5′	CTTTGCACGATGTCTCAACCA	3′
Link-GEFII REV	5′	TTTCCCGTGGAGCAGGAA	3′
CTEN	FAM 5′	CCGCCGCCTAATATGCAACATTAGGG	3′TAMRA
CTEN FOR	5′	CGAGTATTCCAAAGCTGGTATCG	3′
CTEN REV	5′	ATCACAGAGAGATGGCCCTTATCT	3′
NAP4	FAM 5′	TCCGCCTCAGTCGCCTCTTTCG	3′TAMRA
NAP4 FOR	5′	TCGGAAGGGCTCCTTCAAA	3′
NAP4 REV	5′	CACCGTTGCAGCTCTTGGT	3′
MRLP45	FAM 5′	CTCCCATTCCCCTCATGCTATAAAAAGAACTACC	3′TAMRA
MRLP45 FOR	5′	GGCTGCTGGAAGCTTTGAAG	3′
MRLP45 REV	5′	TGAGCAGGATGGGAGAGAACA	3′
TCF2	FAM 5′	CAAAAGCTGGCCATGGACGCCT	3′TAMRA
TCF2 FOR	5′	GCAGGAAGGAGGAGGCATTC	3′
TCF2 REV	5′	CAGGCTGTGAGTCTGGTTGGA	3′
ROK1	FAM 5′	CAGCTGGCTTCCATTTTCCTGGCCT	3′TAMRA
ROK1 FOR	5′	TGGCAAAACTGGGTTCAGAGA	3′

Table 5 (continued)

PRIMER		SEQUENCE	
ROK1 REV	5′	TCGGACCTTGTGGGATGTG	3′
KRT1	FAM 5′	CCGCCGCCTAATATGCAACATTAGGG	3′TAMRA
KRT1 FOR	5′	CGAGTATTCCAAAGCTGGTATCG	3′
KRT1 REV	5′	ATCACAGAGAGATGGCCCTTATCT	3′
KRT5	FAM 5′	CCGCCGCCTAATATGCAACATTAGGG	3′TAMRA
KRT5 FOR	5′	CGAGTATTCCAAAGCTGGTATCG	3′
KRT5 REV	5′	ATCACAGAGAGATGGCCCTTATCT	3′
KRT8	FAM 5′	CCGCCGCCTAATATGCAACATTAGGG	3′TAMRA
KRT8 FOR	5′	CGAGTATTCCAAAGCTGGTATCG	3′
KRT8 REV	5′	ATCACAGAGAGATGGCCCTTATCT	3′
KRT9	FAM 5′	CCGCCGCCTAATATGCAACATTAGGG	3′TAMRA
KRT9 FOR	5′	CGAGTATTCCAAAGCTGGTATCG	3′
KRT9 REV	5′	ATCACAGAGAGATGGCCCTTATCT	3′
KRT10-2	FAM 5′	CCGCCGCCTAATATGCAACATTAGGG	3′TAMRA
KRT10-2 FOR	5′	CGAGTATTCCAAAGCTGGTATCG	3′
KRT10-2 REV	5′	ATCACAGAGAGATGGCCCTTATCT	3′
KRT14	FAM 5′	CCGCCGCCTAATATGCAACATTAGGG	3′TAMRA
KRT14 FOR	5′	CGAGTATTCCAAAGCTGGTATCG	3′
KRT14 REV	5′	ATCACAGAGAGATGGCCCTTATCT	3′
KRT18	FAM 5′	CCGCCGCCTAATATGCAACATTAGGG	3′TAMRA
KRT18 FOR	5′	CGAGTATTCCAAAGCTGGTATCG	3′
KRT18 REV	5′	ATCACAGAGAGATGGCCCTTATCT	3′
KRT19	FAM 5′	CCGCCGCCTAATATGCAACATTAGGG	3′TAMRA
KRT19 FOR	5′	CGAGTATTCCAAAGCTGGTATCG	3′
KRT19 REV	5′	ATCACAGAGAGATGGCCCTTATCT	3′
KRT6a/b	FAM 5′	CCGCCGCCTAATATGCAACATTAGGG	3′TAMRA
KRT6a/b FOR	5′	CGAGTATTCCAAAGCTGGTATCG	3′
KRT6a/b REV	5′	ATCACAGAGAGATGGCCCTTATCT	3′
KRT20	FAM 5′	TGGCGGGAATCCTATTTATCAGACTCTGTAATTG A	3′TAMRA

Table 5 (continued)

PRIMER	SEQUENCE
KRT20 FOR	5′ GCAAGAAATCAGCCATAAGAAAGC 3′
KRT20 REV	5′ TTGCAGCTCCTCTGAGTAAAACAT 3′

Table 6

No.	ID	forward	reverse	PCR size (bp)	GB_ID
1	D17S946	ACAGTCTATCAAGCAGAAAAATCCT	TGCCGTGCCAGAGAGA	128-142	Z24029
2	D17S1181	GACAACAGAGCGAGACTCCC	GCCCAGCCTGTCACTTATTC	122	-
3	D17S2026	TGGTCATTCGACAACGAA	CAGCATTGGATGCAATCC	171-318	G05498 X53777
4	D17S838	CTCCAGAATCCAGACCATGA	AGGACAGTGTGTAGCCCTTC	71-103	Z51080
5	D17S250	GGAAGAATCAAATAGACAAT	GCTGGCCATATATATATTTAAACC	151	-
6	D17S1818	CATAGGTATGTTCAGAAATGTGA	TGCCTACTGGAAACCAGA	119-151	Z52895
7	D17S614	AAGGGGAAGGGGCTTTCAAAGCT	NGGAGGTTGCAGTGAGCCAAGAT	136	L29873
8	D17S2019	CAAAAGCTTATGATGCTCAAACC	TTGTTTCCCTTTGACTTTCTGA	151-152	G07286 Z39013
9	D17S608	TAGGTTCACCTCTCATTTTCTTCAG	GTCTGGGTCTTTATGGNGCTTGTG	136	L29870
10	D17S1655	CGGACCAGAGTGTTCCATGG	GCATACAGCACCCTCTACCT	240	-
11	D17S2147	AGGGGAGAATAAATAAAATCTGTGG	CAGGAGTGAGACACTCTCCATG	138	G15195
12	D17S754	TGGATTCACTGACTCAGCCTGC	GCGTGTCTGTCTCCATGTGTGC	145	-
13	D17S1814	TCCCCAATGACGGTGATG	CTGGAGGTTGGCTTGTGGAT	150-166	Z52854
14	D17S2007	GGTCCCACGAATTTGCTG	CCACCCAGAAAAACAGGAGA	102-103	G07073 X03438
15	D17S1246	TCGATCTCCTGACCTTGTGA	TTGTCACCCCATTGCCTTTC	115	-
16	D17S1979	CCTTGGATAGATTCAGCTCCC	CTTGTCCCTTCTCAATCCTCC	199	G11172 X55068
17	D17S1984	TTAAGCAAGGTTTTAATTAAGCTGC	GATTACAGTGCTCCCTCTCCC	134	G14779 T50487
18	D17S1984	GGTTTTAATTAAGCTGCATGGC	GATTACAGTGCTCCCTCTCCC	126	G11580 T50487
19	D17S1867	AGTTTGACACTGAGGCTTTG	TTTAGACTTGGTAACTGCCG	94	Z51301
20	D17S1788	TGCAGATGCCTAAGAACTTTTCAG	GCCATGATCTCCCAAAGCC	156-168	Z52160
21	D17S1836	TCGAGGTTATGGTGAGCC	AAACTGTGTGTGTCAAAGGATACT	167-173	Z53182
22	D17S1787	GCTGATCTGAAGCCAATGA	TACATGAAGGCATGGTCTG	239-251	Z52130
23	D17S1660	CTAATATAATCCTGGGCACATGG	GCTGCGGACCAGACAGAT	201	G06069
24	D17S2154	GATAAAAACAAGCACTGGCTCC	CCCACGGCTTTCTTGATCTA	137	G15440
25	D17S1955	TGTAATGTAAGCCCCATGAGG	CACTCAACTCAACAGTCTAAAGGTG	180	G11900
26	D17S2098	GTGAGTTCAAGCATAGTAATTATCC	ATTCAGCCTCAGTTCACTGCTTC	181	G13994
27	D17S518	GATCCAGTGGAGACTCAGAG	TAGTCTCTGGGACACCCAGA	88 - 100	X60690
28	D17S1851	ATTCCTGAGTGTCTACCCTGTTGAG	ACTGACTGCGCCACTGC	237 - 253	Z53675
29	D11S4358	TCGAGAAGGACAAAATCACC	GAACAGGGTTAGTCCATTCG	58	-

Table 6 (continued)

No.	ID	forward	reverse	PCR size (bp)	GB ID
30	D17S964	GTTCTTTCCTCTTGTGGGG	AGTCAGCTGAGATTGTGCC	224	L36695
31	D19S1091	CAAGCCAAGACATCCCAGTT	CCCCACACACAGCTCATATG	238	G14589
32	D17S1179	TTTTCTCTCTCATTCCATTGGG	GCAACAGAGGGAGACTCCAA	113 - 125	-
33	D10S2160	TCCCATCCCGTAAGACCTC	TATGGAGTACCTACTCTATGCCAGG	349	G06592
34	D17S1230	ATTCAAAGCTGGATCCCTTT	AGCTGTGACAAATGCCTGTA	108	L32949
35	D17S1338	TCACCTGAGATTGGGAGACC	AAGATGGGGCAGGAATGG	178 - 200	-
36	D17S2011	TCACTGTCCTCCAAGCCAG	AAACACCACACTCTCCCCTG	115	G07143
37	D17S1237	TTCTTGGGCTTCCCGTAGCC	GGGGCAGACGACTTCTCCTT	186	L32947
38	D17S2038	GGGGATACAACCTTTAAAGTTCC	ATTCACCTAATGAGGATTCTTCTTT	228	G6219
39	D17S2091	GCTGAAATAGCCATCTTGAGCTAC	TCCGCATCCTTTTTAAGAGGCAC	157	G13941
40	D17S649	CTTTCACTCTTTCAGCTGAAGAGG	TGACGTGCTATTTCCTGTTTTGTCT	146	L36685
41	D17S1190	GTTTGTTGCTATGCCTGC	CAACACACTACCCCAGGA	122	L18197
42	M87506	ACTCCTCATCTGTAGGGTCT	GAGTCCGCTACCTGAGTGCT	102 - 120	m87506

CLAIMS

1. A method for the prediction of response to cancer treatment, by the detection of at least . markers characterized in that the markers are genes and fragments thereof or genomi nucleic acid sequences that are located on one chromosomal region which is altered i malignant neoplasia.

2. The method of claim 1 wherein the treatment is an antibody treatment, antihormona treatment, anti-growth factor treatment, taxol based treatment, anthracyclin base treatment and platinum salt based treatment.

3. The method of claim 1 wherein the treatment includes Herceptin™, trastuzumab or 2C antibodies.

4. The method of claim 1 characterized in that the markers are:

   a) genes that are located on one or more chromosomal region(s) which is/are altered in malignant neoplasia; and

   b)

      i) receptor and ligand; or
      ii) members of the same signal transduction pathway; or
      iii) members of synergistic signal transduction pathways; or
      iv) members of antagonistic signal transduction pathways; or
      v) transcription factor and transcription factor binding site; or
      vi) integral parts of heteromeric complexes

5. The method of claim 1 or 2 wherein the malignant neoplasia is breast cancer, ovarian cancer, gastric cancer, colon cancer, esophageal cancer, mesenchymal cancer, bladder cancer or non-small cell lung cancer.

6. The method of any of claims 1 to 5 wherein at least one chromosomal region is defined as the cytogenetic region: 1p13, 1q32, 3p21-p24, 5p13-p14, 8q23-q24, 11q13, 12q13,17q12-q24, 17q11.2-21.3 or 20q13.

7. A method for the prediction, diagnosis or prognosis of malignant neoplasia by the detection of at least one marker characterized in that the marker is selected from:

a) a polynucleotide or polynucleotide analog comprising at least one of the sequences of SEQ ID NO: **319 to 389;**

b) a polynucleotide or polynucleotide analog which hybridizes under stringent conditions to a polynucleotide specified in (a) and encodes a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

c) a polynucleotide or polynucleotide analog the sequence of which deviates from the polynucleotide specified in (a) and (c) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

d) a polynucleotide or polynucleotide analog which represents a specific fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d)

e) a purified polypeptide encoded by a polynucleotide or polynucleotide analog sequence specified in (a) to (e)

f) e purified polypeptide comprising at least one of the sequences of SEQ ID NO: 397 - 467;

Are detected.

8. The method according to any of claims 1 to 6 wherein the markers are selected from:

a) a polynucleotide or polynucleotide analog comprising at least one of the sequences of SEQ ID NO: 2 to 6, 8, 9, 11 to 16, 18, 19, 21 to 26, 53 to 76 or 315 to 389

b) a polynucleotide or polynucleotide analog which hybridizes under stringent conditions to a polynucleotide specified in (a) and encodes a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

c) a polynucleotide or polynucleotide analog the sequence of which deviates from the polynucleotide specified in (a) and (b) due to the generation of the genetic code encoding a polypeptide exhibiting the same biological function as specified for the respective sequence in Table 2 or 3

d) a polynucleotide or polynucleotide analog which represents a specific fragmer derivative or allelic variation of a polynucleotide sequence specified in (a) to (c)

e) a purified polypeptide encoded by a polynucleotide sequence or polynucleotic analog specified in (a) to (d)

f) A purified polypeptide comprising at least one of the sequences of SEQ ID NO: 2 to 52 or 76 to 98 or 393 to 467

are detected.

9. A diagnostic kit for conducting the method of claims 1 to 8.

METHODS AND COMPOSITIONS FOR THE PREDICTION, DIAGNOSIS, PROGNOSIS, PREVENTION AND TREATMENT OF MALIGNANT NEOPLASIA

ABSTRACT OF THE DISCLOSURE

The invention provides novel compositions, methods and uses, for the prediction., diagnosis, prognosis, prevention and treatment of malignant neoplasia and breast cancer in particular. Genes that are differentially expressed in breast tissue of breast cancer patients versus those of normal people are disclosed.

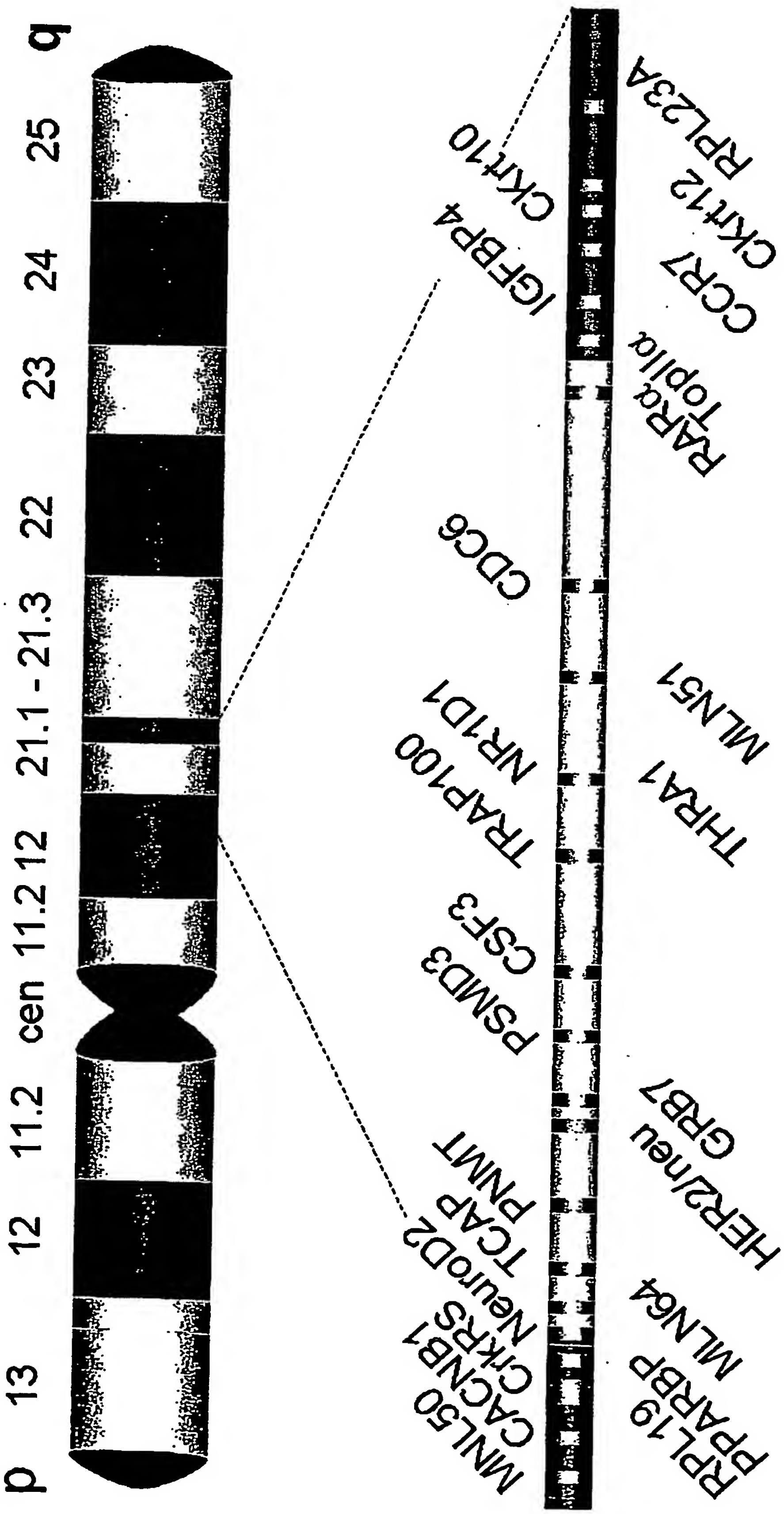
p
13
12
11.2
cen
11.2
12
21.1 - 21.3
21.1 - 21.3
22
23
24
25
q
MNL50
CACNB1
CrkRS
NeuroD2
TCAP
PNMT
PSMD3
CSF3
TRAP100
NR1D1
CDC6
IGFBP4
CKrt10
RPL19
PPARBP
MLN64
HER2/neu
GRB7
THRA1
MLN51
RARα
TopIIα
CCR7
CKrt12
RPL23A

Figure 1

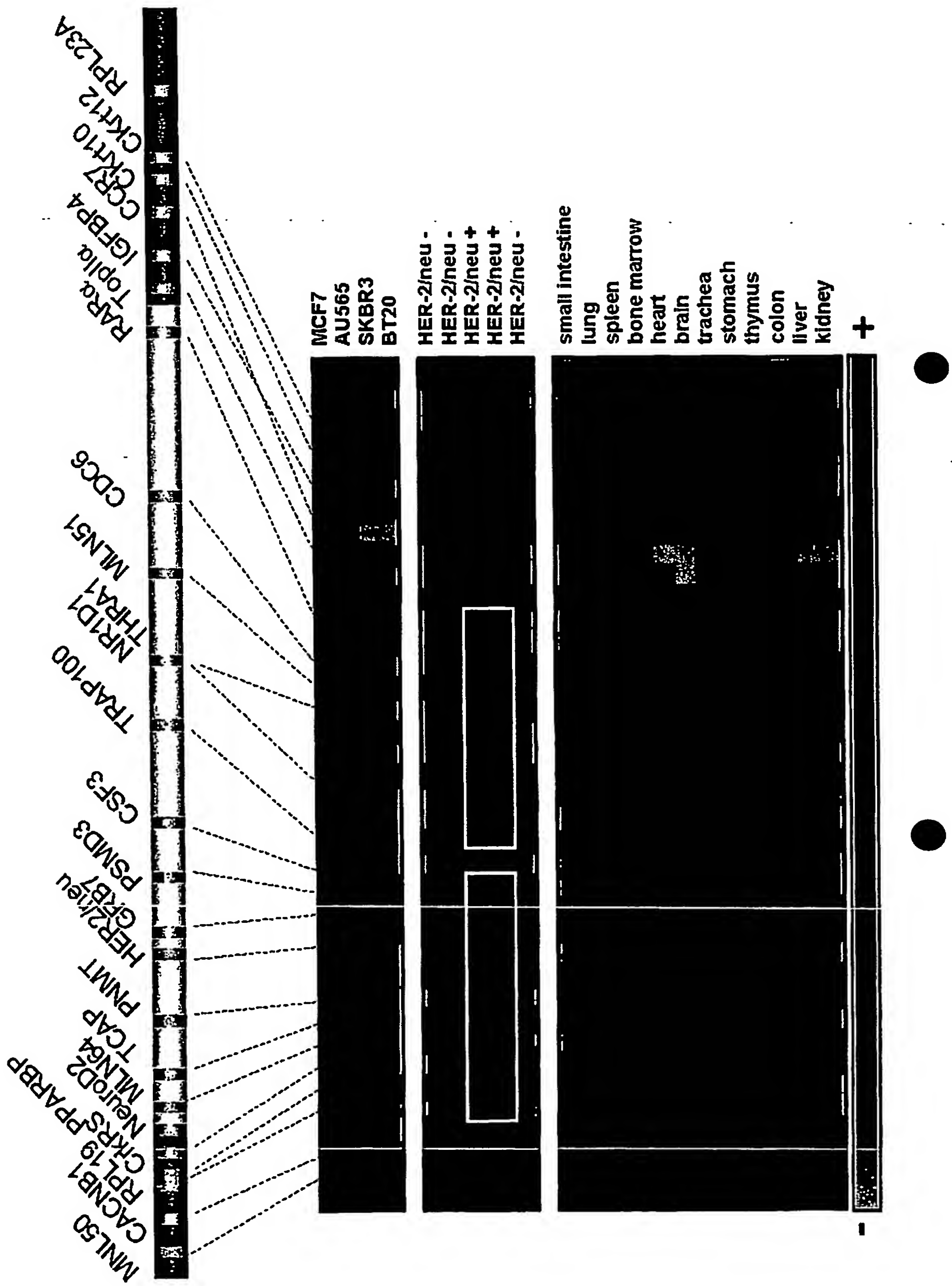
RPL23A
CKt12
CKt10
CCR7
IGFBP4
Toplla
RARα
CDC6
MLN51
THRA1
NR1D1
TRAP100
CSF3
PSMD3
GRB7
HER2/neu
PNMT
TCAP
MLN64
NeuroD2
CrkRS
PPARBP
RPL19
CACNB1
MNL50
MCF7
AU565
SKBR3
BT20
HER-2/neu -
HER-2/neu -
HER-2/neu +
HER-2/neu +
HER-2/neu -
small intestine
lung
spleen
bone marrow
heart
brain
trachea
stomach
thymus
colon
liver
kidney
+
-

Figure 2

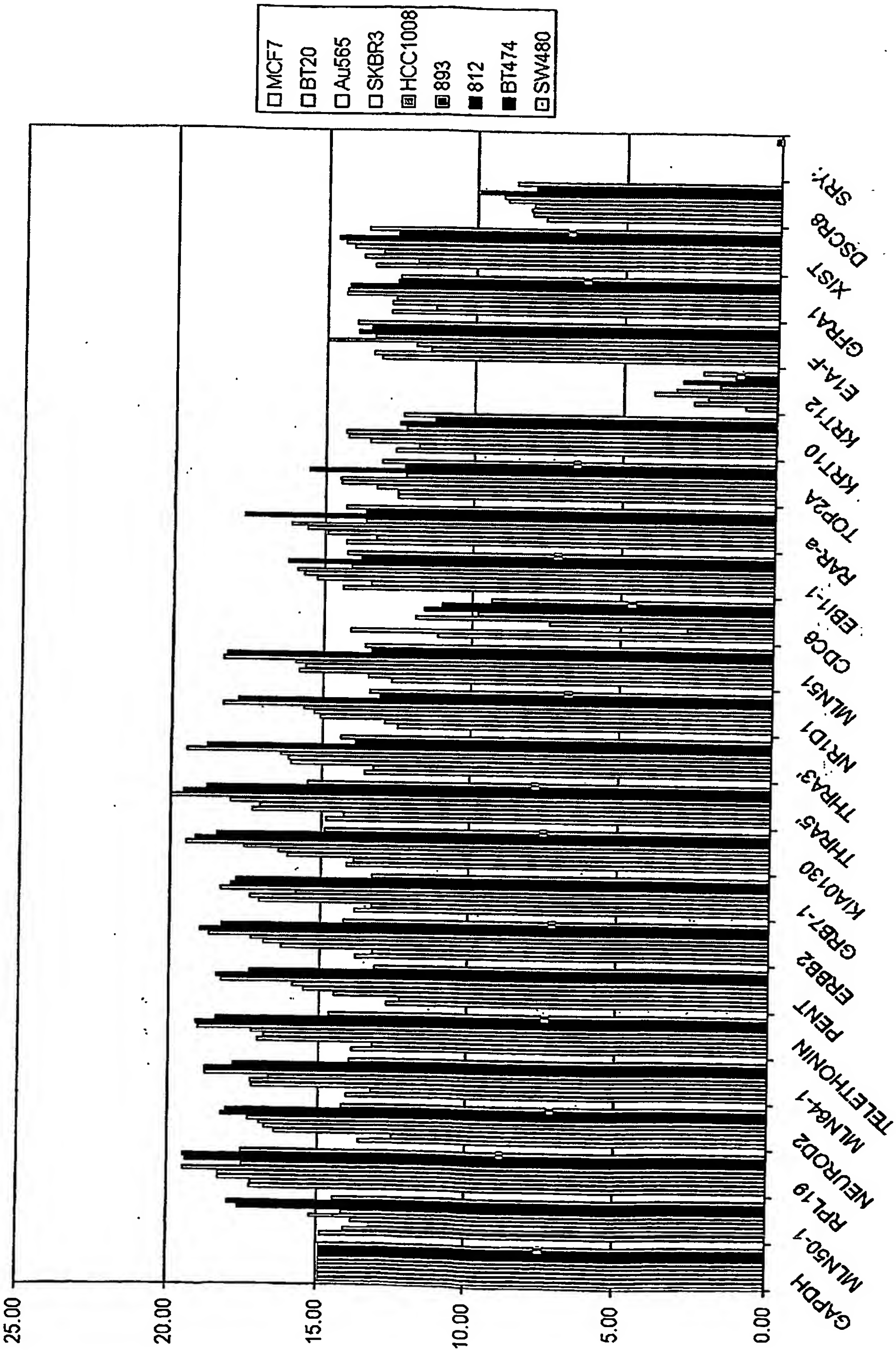
MCF7
BT20
Au565
SKBR3
HCC1008
893
812
BT474
SW480
25.00
20.00
15.00
10.00
5.00
0.00
GAPDH
MLN50-1
RPL19
NEURODZ
MLN64-1
TELETHONIN
PENT
ERBB2
GRB7-1
KIAA0130
THRA5
THRA3
NR1D1
MLN51
CDC6
EBI1-1
RAR-a
TOP2A
KRT10
KRT12
E1A-F
GFRA1
XIST
DSCR8
SRY

Figure 3

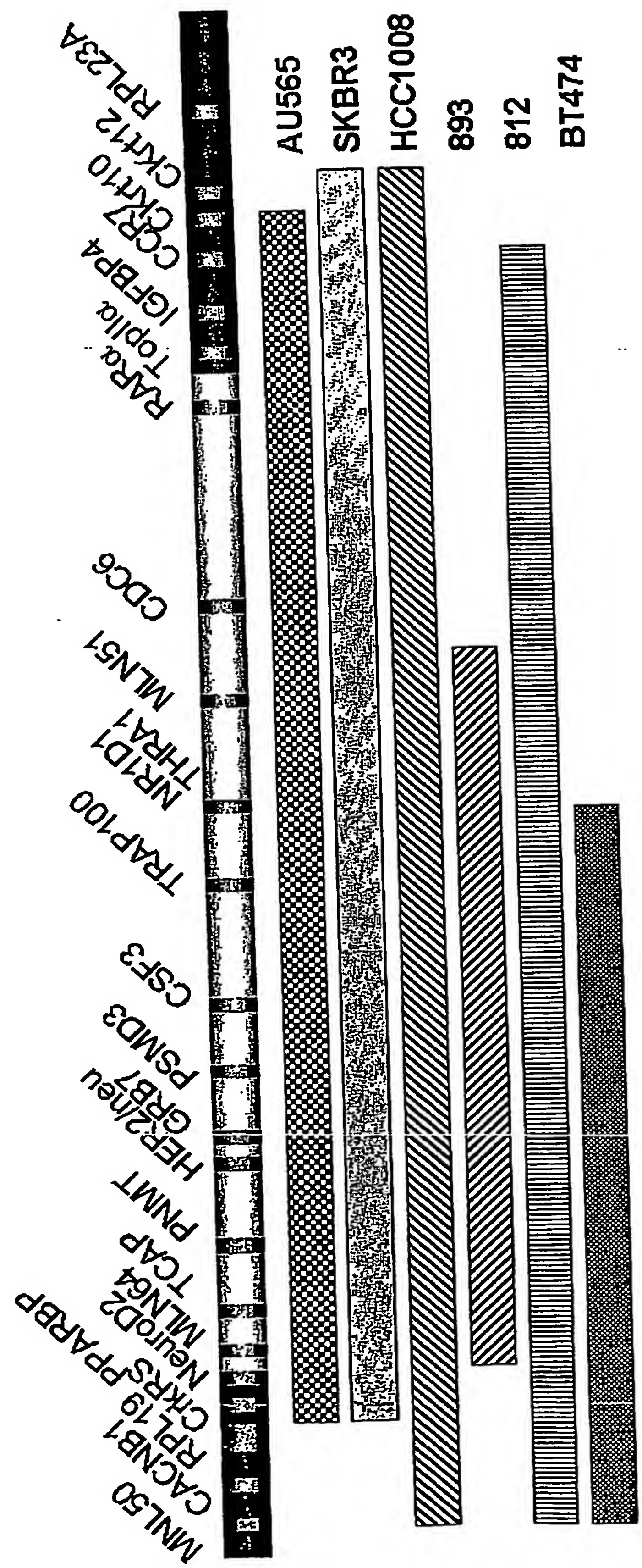
MNL50
CACNB1
RPL19
CkRS
PPARBP
NeuroD2
MLN64
TCAP
PNMT
HER2/neu
GRB7
PSMD3
CSF3
TRAP100
NR1D1
THRA1
MLN51
CDC6
RARα
Topllα
IGFBP4
CCR7
Ckt10
Ckt12
RPL23A
AU565
SKBR3
HCC1008
893
812
BT474

Figure 4

EPO - Munich
9
28. Okt. 2003

SEQUENCE LISTING

<110> Bayer AG

<120> METHODS AND COMPOSITIONS FOR THE PREDICTION, DIAGNOSIS, PROGNOSIS, PREVENTION AND TREATMENT OF MALIGNANT NEOPLASIA

<130> BHC031051

<160> 512

<170> PatentIn version 3.1

<210> 1

<211> 3846

<212> DNA

<213> Homo sapiens

<400> 1

```
gcctcccgcc agctcgcctc ggggaacagg acgcgcgtga gctcaggcgt ccccgcccca      60
gcttttctcg gaaccatgaa ccccaactgc gcccggtgcg gcaagatcgt gtatcccacg     120
gagaaggtga actgtctgga taagttctgg cataaagcat gcttccattg cgagacctgc     180
aagatgacac tgaacatgaa gaactacaag ggctacgaga agaagcccta ctgcaacgca     240
cactacccca agcagtcctt caccatggtg gcggacaccc cggaaaacct tcgcctcaag     300
caacagagtg agctccagag tcaggtgcgc tacaaggagg agtttgagaa gaacaagggc     360
aaaggtttca gcgtagtggc agacacgccc gagctccaga gaatcaagaa gacccaggac     420
cagatcagta atataaaata ccatgaggag tttgagaaga gccgcatggg ccctagcggg     480
ggcgagggca tggagccaga gcgtcgggat tcacaggacg gcagcagcta ccggcggccc     540
ctggagcagc agcagcctca ccacatcccg accagtgccc cggtttacca gcagccccag     600
cagcagccgg tggcccagtc ctatggtggc tacaaggagc ctgcagcccc agtctccata     660
cagcgcagcg ccccaggtgg tggcgggaag cggtaccgcg cggtgtatga ctacagcgcc     720
gccgacgagg acgaggtctc cttccaggac ggggacacca tcgtcaacgt gcagcagatc     780
gacgacggct ggatgtacgg gacggtggag cgcaccggcg acacggggat gctgccggcc     840
aactacgtgg aggccatctg aacccggagc gcccccatct gtcttcagca cattccacgg     900
catcgcatcc gtcctgggcg tgagccgtcc attcttcagt gtctctgttt tttaaaacct     960
gcgacagctt gtgattccta cccctcttcc agcttctttt gccaactgaa gccttcttct    1020
gccacttctg cgggctccct cctctggcag gcttcccccg tgatcgactt cttggttttc    1080
tctctggatg gaacgggtat gggcctctct gggggaggca gggctggaat gggagacctg    1140
ttggcctgtg ggcctcacct gcccctctgt tctctcccct cacatcctcc tgcccagctc    1200
ctcacatacc cacacattcc agggctgggg tgagcctgac tgccaggacc ccaggtcagg    1260
ggctccctac attccccaga gtgggatcca cttcttggtt cctgggatgg cgatggggac    1320
tctgccgctg tgtagggacc agtgggatgg gctctacctc tctttctcaa agagggggct    1380
ctgcccacct ggggtctctc tccctacctc cctcctcagg ggcaacaaca ggagaatggg    1440
gttcctgctg tggggcgaat tcatcccctc cccgcgcgtt ccttcgcaca ctgtgatttt    1500
gccctcctgc ccacgcagac ctgcagcggg caaagagctc ccgaggaagc acagcttggg    1560
tcaggttctt gcctttctta attttaggga cagctaccgg aaggagggga acaaggagtt    1620
ctcttccgca gcccctttcc ccacgcccac ccccagtctc cagggaccct tgcctgcctc    1680
ctaggctgga agccatggtc ccgaagtgta gggcaagggt gcctcaggac cttttggtct    1740
tcagcctccc tcagccccca ggatctgggt taggtggccg ctcctccctg ctcctcatgg    1800
```

gaagatgtct cagagccttc catgacctcc cctccccagc ccaatgccaa gtggacttgg 1860
agctgcacaa agtcagcagg gaccactaaa tctccaagac ctggtgtgcg gaggcaggag 1920
catgtatgtc tgcaggtgtc tgacacgcaa gtgtgtgagt gtgagtgtga gagatggggc 1980
gggggtgtgt ctgtaggtgt ctctgggcct gtgtgtgggt ggggttatgt gagggtatga 2040
agagctgtct tcccctgaga gtttcctcag aacccacagt gagaggggag ggctcctggg 2100
gcagagaagt tccttaggtt ttctttggaa tgaaattcct ccttcccccc atctctgagt 2160
ggaggaagcc caccaatctg ccctttgcag tgtgtcaggg tggaaggtaa gaggttggtg 2220
tggagttggg gctgccatag ggtctgcagc ctgctggggc taagcggtgg aggaaggctc 2280
tgtcactcca ggcatatgtt tccccatctc tgtctggggc tacagaatag ggtggcagaa 2340
gtgtcaccct gtgggtgtct ccctcggggg ctcttcccct agacctcccc ctcacttaca 2400
taaagctccc ttgaagcaag aaagagggtc ccagggctgc aaaactggaa gcacagcctc 2460
ggggatgggg agggaaagac ggtgctatat ccagttcctg ctctctgctc atgggtggct 2520
gtgacaaccc tggcctcact tgattcatct ctggttttct tgccaccctc tgggagtccc 2580
catcccattt tcatcctgag cccaaccagg ccctgccatt ggcctcttgt cccttggcac 2640
acttgtaccc acaggtgagg ggcaggacct gaaggtattg gcctgttcaa caatcagtca 2700
tcatgggtgt ttttgtcaac tgcttgttaa ttgatttggg gatgtttgcc ccgaatgaga 2760
ggttgaggaa aagactgtgg gtggggaggc cctgcctgac ccatcccttt tcctttctgg 2820
ccccagccta ggtggaggca agtggaatat cttatattgg gcgatttggg ggctcgggga 2880
ggcagagaat ctcttgggag tcttgggtgg cgctggtgca ttctgtttcc tcttgatctc 2940
aaagcacaat gtggatttgg ggaccaaagg tcagggacac atccccttag aggacctgag 3000
tttgggagag tggtgagtgg aagggaggag cagcaagaag cagcctgttt tcactcagct 3060
taattctcct tcccagataa ggcaagccag tcatggaatc ttgctgcagg ccctccctct 3120
actcttcctg tcctaaaaat aggggccgtt ttcttacaca cccccagaga gaggagggac 3180
tgtcacactg gtgctgagtg accgggggct gctgggcgtc tgttctttac caaaaccatc 3240
catccctaga agagcacaga gccctgaggg gctgggctgg gctgggctga gcccctggtc 3300
ttctctacag ttcacagagg tctttcagct catttaatcc caggaaagag gcatcaaagc 3360
tagaatgtga atataacttt tgtgggccaa tactaagaat aacaagaagc ccagtggtga 3420
ggaaagtgcg ttctcccagc actgcctcct gttttctccc tctcatgtcc ctccagggaa 3480
aatgacttta ttgcttaatt tctgcctttc ccccctcaca catgcacttt tgggcctttt 3540
tttatagctg gaaaaaacaa aataccaccc tacaaacctg tatttaaaaa gaaacagaaa 3600
tgaccacgtg aaatttgcct ctgtccaaac atttcatccg tgtgtatgtg tatgtgtgtg 3660
agtgtgtgaa gccgccagtt catcttttta tatggggttg ttgtctcatt ttggtctgtt 3720
ttggtcccct ccctcgtggg cttgtgctcg ggatcaaacc tttctggcct gttatgattc 3780
tgaacatttg acttgaacca caagtgaatc tttctcctgg tgactcaaat aaaagtataa 3840
tttta 3846

<210> 2

<211> 1711

<212> DNA

<213> Homo sapiens

<400> 2

gagggaaggc aggaaggagg cagccgaagg ccgagctggg tggctggacc gggtgctggc 60
tgcgcgcgct gctttcggct cccacggcct ctcccatgcg ctgagggagc ccggctgcgg 120
gccggcggcg ggagggaggg ctcctctcca tggtccagaa gaccagcatg tcccggggcc 180
cttacccacc ctcccaggag atccccatgg aggtcttcga ccccagcccg cagggcaaat 240
acagcaagag gaaagggcga ttcaaacggt cagatgggag cacgtcctcg gataccacat 300
ccaacagctt tgtccgccag ggctcagcgg agtcctacac cagccgtcca tcagactctg 360
atgtatctct ggaggaggac cgggaagcct taaggaagga agcagagcgc caggcattag 420
cgcagctcga gaaggccaag accaagccag tggcatttgc tgtgcggaca aatgttggct 480
acaatccgtc tccaggggat gaggtgcctg tgcagggagt ggccatcacc ttcgagccca 540
aagacttcct gcacatcaag gagaaataca ataatgactg gtggatcggg cggctggtga 600
aggagggctg tgaggttggc ttcattccca gccccgtcaa actggacagc cttcgcctgc 660
tgcaggaaca gaagctgcgc cagaaccgcc tcggctccag caaatcaggc gataactcca 720
gttccagtct gggagatgtg gtgactggca cccgccgccc cacacccccct gccagtgcca 780
aacagaagca gaagtcgaca gagcatgtgc ccccctatga cgtggtgcct tccatgaggc 840
ccatcatcct ggtgggaccg tcgctcaagg gctacgaggt tacagacatg atgcagaaag 900
ctttatttga cttcttgaag catcggtttg atggcaggat ctccatcact cgtgtgacgg 960
cagatatttc cctggctaag cgctcagttc tcaacaaccc cagcaaacac atcatcattg 1020
agcgctccaa cacacgctcc agcctggctg aggtgcagag tgaaatcgag cgaatcttcg 1080
agctggcccg gacccttcag ttggtcgctc tggatgctga caccatcaat cacccagccc 1140
agctgtccaa gacctcgctg gcccccatca ttgtttacat caagatcacc tctcccaagg 1200

tacttcaaag gctcatcaag tcccgaggaa agtctcagtc caaacacctc aatgtccaaa 1260
tagcggcctc ggaaaagctg gcacagtgcc cccctgaaat gtttgacatc atcctggatg 1320
agaaccaatt ggaggatgcc tgcgagcatc tggcggagta cttggaagcc tattggaagg 1380
ccacacaccc gcccagcagc acgccaccca atccgctgct gaaccgcacc atggctaccg 1440
cagccctgcg ccgtagccct gcccctgtct ccaacctcca ggtacaggtg ctcacctcgc 1500
tcaggagaaa cctcggcttc tggggcgggc tggagtcctc acagcggggc agtgtggtgc 1560
cccaggagca ggaacatgcc atgtagtggg cgccctgccc gtcttccctc ctgctctggg 1620
gtcggaactg gagtgcaggg aacatggagg aggaagggaa gagctttatt ttgtaaaaaa 1680
ataagatgag cggcaaaaaa aaaaaaaaaa a 1711

<210> 3

<211> 698

<212> DNA

<213> Homo sapiens

<400> 3

ttttcctttc gctgctgcgg ccgcagccat gagtatgctc aggcttcaga agaggctcgc 60
ctctagtgtc ctccgctgtg gcaagaagaa ggtctggtta gaccccaatg agaccaatga 120
aatcgccaat gccaactccc gtcagcagat ccggaagctc atcaaagatg ggctgatcat 180
ccgcaagcct gtgacggtcc attcccgggc tcgatgccgg aaaaacacct tggcccgccg 240
gaagggcagg cacatgggca taggtaagcg gaagggtaca gccaatgccc gaatgccaga 300
gaaggtcaca tggatgagga gaatgaggat tttgcgccgg ctgctcagaa gataccgtga 360
atctaagaag atcgatcgcc acatgtatca cagcctgtac ctgaaggtga aggggaatgt 420
gttcaaaaac aagcggattc tcatggaaca catccacaag ctgaaggcag acaaggcccg 480
caagaagctc ctggctgacc aggctgaggc ccgcaggtct aagaccaagg aagcacgcaa 540
gcgccgtgaa gagcgcctcc aggccaagaa ggaggagatc atcaagactt tatccaagga 600
ggaagagacc aagaaataaa acctcccact ttgtctgtac atactggcct ctgtgattac 660
atagatcagc cattaaaata aaacaagcct taatctgc 698

<210> 4

<211> 5810

<212> DNA

<213> Homo sapiens

<400> 4

gggaagatgg cggcggcctc gagcaccctc ctcttcttgc cgccggggac ttcagattga 60
tccttcccgg gaagagtagg gactgctggt gccctgcgtc ccgggatccc gagccaactt 120
gtttcctccg ttagtggtgg ggaagggctt atcctttttgt ggcggatcta gcttctcctc 180
gccttcagga tgaaagctca gggggaaac cgaggagtca gaaaagctga gtaagatgag 240
ttctctcctg gaacggctcc atgcaaaatt taaccaaaat agaccctgga gtgaaaccat 300
taagcttgtg cgtcaagtca tggagaagag ggttgtgatg agttctggag ggcatcaaca 360
tttggtcagc tgtttggaga cattgcagaa ggctctcaaa gtaacatctt taccagcaat 420
gactgatcgt ttggagtcca tagcaggaca gaatggactg ggctctcatc tcagtgccag 480
tggcactgaa tgttacatca cgtcagatat gttctatgtg gaagtgcagt tagatcctgc 540
aggacagctt tgtgatgtaa aagtggctca ccatggggag aatcctgtga gctgtccgga 600
gcttgtacag cagctaaggg aaaaaaattc tgatgaattt tctaagcacc ttaagggcct 660
tgttaatctg tataaccttc caggggacaa caaactgaag actaaaatgt acttggctct 720
ccaatcctta gaacaagatc tttctaaaat ggcaattatg tactggaaag caactaatgc 780
tggtcccttg gataagattc ttcatggaag tgttggctat ctcacaccaa ggagtggggg 840
tcatttaatg aacctgaagt actatgtctc tccttctgac ctactggatg acaagactgc 900
atctcccatc attttgcatg agaataatgt ttctcgatct ttgggcatga atgcatcagt 960
gacaattgaa ggaacatctg ctgtgtacaa actcccaatt gcaccattaa ttatggggtc 1020
acatccagtt gacaataaat ggacccctc cttctcctca atcaccagtg ccaacagtgt 1080
tgatcttcct gcctgtttct tcttgaaatt tccccagcca atcccagtat ctagagcatt 1140
tgttcagaaa ctgcagaact gcacaggaat tccattgttt gaaactcaac caacttatgc 1200
accctgtat gaactgatca ctcagtttga gctatcaaag gaccctgacc ccatacctttt 1260
gaatcacaac atgagatttt atgctgctct tcctggtcag cagcactgct atttcctcaa 1320
caaggatgct cctcttccag atggccgaag tctacaggga acccttgtta gcaaaatcac 1380

ctttcagcac cctggccgag ttcctcttat cctaaatctg atcagacacc aagtggccta 1440
taacaccctc attggaagct gtgtcaaaag aactattctg aaagaagatt ctcctgggct 1500
tctccaattt gaagtgtgtc ctctctcaga gtctcgtttc agcgtatctt ttcagcaccc 1560
tgtgaatgac tccctggtgt gtgtggtaat ggatgtgcag ggcttaacac atgtgagctg 1620
taaactctac aaagggctgt cggatgcact gatctgcaca gatgacttca ttgccaaagt 1680
tgttcaaaga tgtatgtcca tccctgtgac gatgagggct attcggagga aagctgaaac 1740
cattcaagcc gacaccccag cactgtccct cattgcagag acagttgaag acatggtgaa 1800
aaagaacctg cccccggcta gcagcccagg gtatggcatg accacaggca acaacccaat 1860
gagtggtacc actacatcaa ccaacacctt tccgggggt cccattgcca ccttgtttaa 1920
tatgagcatg agcatcaaag atcggcatga gtcggtgggc catggggagg acttcagcaa 1980
ggtgtctcag aacccaattc ttaccagttt gttgcaaatc acagggaacg gggggtctac 2040
cattggctcg agtccgaccc ctcctcatca cacgccgcca cctgtctctt cgatggccgg 2100
caacaccaag aaccacccga tgctcatgaa ccttctcaaa gataatcctg cccaggattt 2160
ctcaaccctt tatggaagca gccctttaga aaggcagaac tcctcttccg gctcaccccg 2220
catggaaata tgctcgggga gcaacaagac caagaaaaag aagtcatcaa gattaccacc 2280
tgagaaacca aagcaccaga ctgaagatga ctttcagagg gagctatttt caatggatgt 2340
tgactcacag aaccctatct ttgatgtcaa catgacagct gacacgctgg atacgccaca 2400
catcactcca gctccaagcc agtgtagcac tcccccaaca acttacccac aaccagtacc 2460
tcacccccaa cccagtattc aaaggatggt ccgactatcc agttcagaca gcattggccc 2520
agatgtaact gacatccttt cagacattgc agaagaagct tctaaacttc ccagcactag 2580
tgatgattgc ccagccattg gcacccctct tcgagattct tcaagctctg ggcattctca 2640
gagtaccctg tttgactctg atgtctttca aactaacaat aatgaaaatc catacactga 2700
tccagctgat cttattgcag atgctgctgg aagccccagt agtgactctc ctaccaatca 2760
tttttttcat gatggagtag atttcaatcc tgatttattg aacagccaga gccaaagtgg 2820
ttttggagaa gaatattttg atgaaagcag ccaaagtggg gataatgatg atttcaaagg 2880
atttgcatct caggcactaa atactttggg ggtgccaatg cttggaggtg ataatgggga 2940
gaccaagttt aagggcaata accaagccga cacagttgat ttcagtatta tttcagtagc 3000
cggcaaagct ttagctcctg cagatcttat ggagcatcac agtggtagtc agggtccttt 3060
actgaccact ggggacttag ggaaagaaaa gactcaaaag agggtaaagg aaggcaatgg 3120
caccagtaat agtactctct cggggcccgg attagacagc aaaccaggga agcgcagtcg 3180
gaccccttct aatgatggga aagcaaaga taagcctcca aagcggaaga aggcagacac 3240
tgagggaaag tctccatctc atagttcttc taacagacct tttaccccac ctaccagtac 3300
aggtggatct aaatcgccag gcagtgcagg aagatctcag actcccccag gtgttgccac 3360
accacccatt cccaaaatca ctattcagat tcctaaggga acagtgatgg tgggcaagcc 3420
ttcctctcac agtcagtata ccagcagtgg ttctgtgtct tcctcaggca gcaaaagcca 3480
ccatagccat tcttcctcct cttcctcatc tgcttccacc tcagggaaga tgaaaagcag 3540
taaatcagaa ggttcatcaa gttccaagtt aagtagcagt atgtattcta gccaggggtc 3600
ttctggatct agccagtcca aaaattcatc ccagtctggg gggaagccag gctcctctcc 3660
cataaccaag catggactga gcagtggctc tagcagcacc aagatgaaac ctcaaggaaa 3720
gccatcatca cttatgaatc cttctttaag taaaccaaac atatcccctt ctcattcaag 3780
gccacctgga ggctctgaca agcttgcctc tccaatgaag cctgttcctg gaactcctcc 3840
atcctctaaa gccaagtccc ctatcagttc aggttctggt ggttctcata tgtctggaac 3900
tagttcaagc tctggcatga agtcatcttc agggttagga tcctcaggct cgttgtccca 3960
gaaaactccc ccatcatcta attcctgtac ggcatcttcc tcctcctttt cctcaagtgg 4020
ctcttccatg tcatcctctc agaaccagca tgggagttct aaaggaaaat ctcccagcag 4080
aaacaagaag ccgtccttga cagctgtcat agataaactg aagcatgggg ttgtcaccag 4140
tggccctggg ggtgaagacc cactggacgg ccagatgggg gtgagcacaa attcttccag 4200
ccatcctatg tcctccaaac ataacatgtc aggaggagag tttcagggca agcgtgagaa 4260
aagtgataaa gacaaatcaa aggtttccac ctccgggagt tcagtggatt cttctaagaa 4320
gacctcagag tcaaaaaatg tggggagcac aggtgtggca aaaattatca tcagtaagca 4380
tgatggaggc tcccctagca ttaaagccaa agtgactttg cagaaacctg gggaaagtag 4440
tggagaaggg cttaggcctc aaatggcttc ttctaaaaac tatggctctc cactcatcag 4500
tggttccact ccaaagcatg agcgtggctc tcccagccat agtaagtcac cagcatatac 4560
cccccagaat ctggacagtg aaagtgagtc aggctcctcc atagcagaga aatcttatca 4620
gaatagtccc agctcagacg atggtatccg accacttcca gaatacagca cagagaaaca 4680
taagaagcac aaaaaggaaa agaagaaagt aaaagacaaa gatagggacc gagaccggga 4740
caaagaccga gacaagaaaa aatctcatag catcaagcca gagagttggt ccaaatcacc 4800
catctcttca gaccagtcct tgtctatgac aagtaacaca atcttatctg cagacagacc 4860
ctcaaggctc agcccagact ttatgattgg ggaggaagat gatgatctta tggatgtggc 4920
cctgattggg aattaggaac cttatttcct aaaagaaaca gggccagagg aaaaaaaact 4980
attgataagt ttataggcaa accaccataa ggggtgagtc agacaggtct gatttggtta 5040
agaatcctaa atggcatggc tttgacatca agctgggtga attagaaagg catatccaga 5100
ccctattaaa gaaaccacag ggtttgattc tggttaccag gaagtcttct ttgttcctgt 5160
gccagaaaga aagttaaaat acttgcttaa gaaagggagg ggggtgggag gggtgtaggg 5220
agagggaagg gagggaaaca gttttgtggg aaatattcat atatattttc ttctcccttt 5280
ttccattttt aggccatgtt ttaaactcat tttagtgcat gtatatgaag ggctgggcag 5340

```
aaaatgaaaa agcaatacat tccttgatgc atttgcatga aggttgttca actttgtttg   5400
aggtagttgt ccgtttgagt catgggcaaa tgaaggactt tggtcatttt ggacacttaa   5460
gtaatgtttg gtgtctgttt cttaggagtg actgggggag ggaagattat tttagctatt   5520
tatttgtaat attttaaccc tttatctgtt tgtttttata cagtgtttcg ttctaaatct   5580
atgaggttta gggttcaaaa tgatggaagg ccgaagagca aggcttatat ggtggtaggg   5640
agcttatagc ttgtgctaat actgtagcat caagcccaag caaattagtc agagcccgcc   5700
tttagagtta aatataatag aaaaaccaaa atgatatttt tatttttagga gggtttaaat   5760
agggttcaga gatcatagga atattaggag ttacctctct gtggaggtat              5810
```

<210> 5

<211> 5515

<212> DNA

<213> Homo sapiens

<400> 5

```
ctttttttccc ttcttcaggt caggggaaag ggaatgccca attcagagag acatgggggc     60
aagaaggacg ggagtggagg agcttctgga actttgcagc cgtcatcggg aggcggcagc    120
tctaacagca gagagcgtca ccgcttggta tcgaagcaca agcggcataa gtccaaacac    180
tccaaagaca tggggttggt gacccccgaa gcagcatccc tgggcacagt tatcaaacct    240
ttggtggagt atgatgatat cagctctgat tccgacacct tctccgatga catggccttc    300
aaactagacc gaagggagaa cgacgaacgt cgtggatcag atcggagcga ccgcctgcac    360
aaacatcgtc accaccagca caggcgttcc cgggacttac taaaagctaa acagaccgaa    420
aaagaaaaaa gccaagaagt ctccagcaag tcgggatcga tgaaggaccg gatatcggga    480
agttcaaagc gttcgaatga ggagactgat gactatggga aggcgcaggt agccaaaagc    540
agcagcaagg aatccaggtc atccaagctc cacaaggaga agaccaggaa agaacgggag    600
ctgaagtctg ggcacaaaga ccggagtaaa agtcatcgaa aaagggaaac acccaaaagt    660
agctccaaac aagatgatag cccctcggga gcttcttatg gccaagatta tgaccttagt    720
ccctcacgat ctcatacctc gagcaattat gactcctaca agaaaagtcc tggaagtacc    780
tcgagaaggc agtcggtcag tccccccttac aaggagcctt cggcctacca gtccagcacc    840
cggtcaccga gcccctacag taggcgacag agatctgtca gtccctatag caggagacgg    900
tcgtccagct acgaaagaag tggctcttac agcgggcgat cgcccagtcc ctatggtcga    960
aggcggtcca gcagcccttt cctgagcaag cggtctctga gtcggagtcc actccccagt   1020
aggaaatcca tgaagtccag aagtagaagt cctgcatatt caagacattc atcttctcat   1080
agtaaaaaga agagatccag ttcacgcagt cgtcattcca gtatctcacc tgtcaggctt   1140
ccacttaatt ccagtctggg agctgaactc agtaggaaaa agaaggaaag agcagctgct   1200
gctgctgcag caaagatgga tggaaaggag tccaagggtt cacctgtatt tttgcctaga   1260
aaagagaaca gttcagtaga ggctaaggat tcaggtttgg agtctaaaaa gttacccaga   1320
agtgtaaaat tggaaaaatc tgccccagat actgaactgg tgaatgtaac acatctaaac   1380
acagaggtaa aaaattcttc agatacaggg aaagtaaagt tggatgagaa ctccgagaag   1440
catcttgtta aagatttgaa agcacaggga acaagagact ctaaacccat agcactgaaa   1500
gaggagattg ttactccaaa ggagacagaa acatcagaaa aggagacccc tccacctctt   1560
cccacaattg cttctccccc acccccctcta ccaactacta cccctccacc tcagacaccc   1620
cctttgccac ctttgcctcc aataccagct cttccacagc aaccacctct gcctccttct   1680
cagccagcat ttagtcaggt tcctgcttcc agtacttcaa ctttgccccc ttctactcac   1740
tcaaagacat ctgctgtgtc ctctcaggca aattctcagc cccctgtaca ggtttctgtg   1800
aagactcaag tatctgtaac agctgctatt ccacacctga aaacttcaac gttgcctcct   1860
ttgcccctcc cacccttatt acctggaggt gatgacatgg atagtccaaa agaaactctt   1920
ccttcaaaac ctgtgaagaa agagaaggaa cagaggacac gtcacttact cacagacctt   1980
cctctccctc cagagctccc tggtggagat ctgtctcccc cagactctcc agaaccaaag   2040
gcaatcacac cacctcagca accatataaa aagagaccaa aaatttgttg tcctcgttat   2100
ggagaaagaa gacaaacaga aagcgactgg gggaaacgct gtgtggacaa gtttgacatt   2160
attgggatta ttggagaagg aacctatggc caagtatata aagccaggga caaagacaca   2220
ggagaactag tggctctgaa gaaggtgaga ctagacaatg agaaagaggg cttcccaatc   2280
acagccattc gtgaaatcaa aatccttcgt cagttaatcc accgaagtgt tgttaacatg   2340
aaggaaattg tcacagataa acaagatgca ctggatttca agaaggacaa aggtgccttt   2400
taccttgtat ttgagtatat ggaccatgac ttaatgggac tgctagaatc tggtttggtg   2460
cacttttctg aggaccatat caagtcgttc atgaaacagc taatggaagg attggaatac   2520
tgtcacaaaa agaatttcct gcatcgggat attaagtgtt ctaacatttt gctgaataac   2580
agtgggcaaa tcaaactagc agatttggga cttgctcggc tctataactc tgaagagagt   2640
cgcccttaca caaacaaagt cattactttg tggtaccgac ctccagaact actgctagga   2700
gaggaacgtt acacaccagc catagatgtt tggagctgtg gatgtattct tggggaacta   2760
```

Note: the position numbers printed at the right of the <400> 5 lines are 60, 120, … 2820, but they appear to be off by one line here. The correct alignment is that the first line ends at 60 and the last at 2820.

ttcacaaaga agcctatttt tcaagccaat ctggaactgg ctcagctaga actgatcagc 2880
cgactttgtg gtagcccttg tccagctgtg tggcctgatg ttatcaaact gccctacttc 2940
aacaccatga aaccgaagaa gcaatatcga aggcgtctac gagaagaatt ctctttcatt 3000
ccttctgcag cacttgattt attggaccac atgctgacac tagatcctag taagcggtgc 3060
acagctgaac agaccctaca gagcgacttc cttaaagatg tcgaactcag caaaatggct 3120
cctccagacc tcccccactg gcaggattgc catgagttgt ggagtaagaa acggcgacgt 3180
cagcgacaaa gtggtgttgt agtcgaagag ccacctccat ccaaaacttc tcgaaaagaa 3240
actacctcag ggacaagtac tgagcctgtg aagaacagca gcccagcacc acctcagcct 3300
gctcctggca aggtggagtc tggggctggg gatgcaatag gccttgctga catcacacaa 3360
cagctgaatc aaagtgaatt ggcagtgtta ttaaacctgc tgcagagcca aaccgacctg 3420
agcatccctc aaatggcaca gctgcttaac atccactcca acccagagat gcagcagcag 3480
ctggaagccc tgaaccaatc catcagtgcc ctgacggaag ctacttccca gcagcaggac 3540
tcagagacca tggccccaga ggagtctttg aaggaagcac cctctgcccc agtgatcctg 3600
ccttcagcag aacagatgac ccttgaagct tcaagcacac cagctgacat gcagaatata 3660
ttggcagttc tcttgagtca gctgatgaaa acccaagagc cagcaggcag tctggaggaa 3720
aacaacagtg acaagaacag tgggccacag ggccccgaa gaactcccac aatgccacag 3780
gaggaggcag cagcatgtcc tcctcacatt cttccaccag agaagaggcc ccctgagccc 3840
cccggacctc caccgccgcc acctccaccc cctctggttg aaggcgatct ttccagcgcc 3900
ccccaggagt tgaacccagc cgtgacagcc gccttgctgc aacttttatc ccagcctgaa 3960
gcagagcctc ctggccacct gccacatgag caccaggcct tgagaccaat ggagtactcc 4020
acccgacccc gtccaaacag gacttatgga aacactgatg ggcctgaaac agggttcagt 4080
gccattgaca ctgatgaacg aaactctggt ccagccttga cagaatcctt ggtccagacc 4140
ctggtgaaga acaggacctt ctcaggctct ctgagccacc ttggggagtc cagcagttac 4200
cagggcacag ggtcagtgca gtttccaggg gaccaggacc tccgttttgc cagggtcccc 4260
ttagcgttac acccggtggt cgggcaacca ttcctgaagg ctgagggaag cagcaattct 4320
gtggtacatg cagagaccaa attgcaaaac tatggggagc tggggccagg aaccactggg 4380
gccagcagct caggagcagg ccttcactgg gggggcccaa ctcagtcttc tgcttatgga 4440
aaactctatc gggggcctac aagagtccca ccaagagggg gaagagggag aggagttcct 4500
tactaaccca gagacttcag tgtcctgaaa gattcctttc ctatccatcc ttccatccag 4560
ttctctgaat ctttaatgaa atcatttgcc agagcgaggt aatcatctgc atttggctac 4620
tgcaaagctg tccgttgtat tccttgctca cttgctacta gcaggcgact taggaaataa 4680
tgatgttggc accagttccc cctggatggg ctatagccag aacatttact tcaactctac 4740
cttagtagat acaagtagag aatatggaga ggatcattac attgaaaagt aaatgtttta 4800
ttagttcatt gcctgcactt actggtcgga agagagaaag aacagtttca gtattgagat 4860
ggctcaggag aggctctttg atttttaaag ttttggggtg gggggttgtg tgtggtttct 4920
ttcttttgaa ttttaattta ggtgttttgg gttttttttcc tttaaagaga atagtgttca 4980
caaaatttga gctgctcttt ggcttttgct ataagggaaa cagagtggcc tggctgattt 5040
gaataaatgt ttctttcctc tccaccatct cacattttgc ttttaagtga acactttttc 5100
cccattgagc atcttgaaca tactttttttt ccaaataaat tactcatcct taaagtttac 5160
tccactttga caaaagatac gcccttctcc ctgcacataa agcaggttgt agaacgtggc 5220
attcttgggc aagtaggtag actttaccca gtctctttcc tttttgctg atgtgtgctc 5280
tctctctctc tttctctctc tctctctctc tctctctctc tctgtctgtc tcgcttgctc 5340
gctctcgctg tttctctctc tttgaggcat ttgtttggaa aaaatcgttg agatgcccaa 5400
gaacctggga taattcttta cttttttgga aataaaggaa aggaaattca aaaaaaaaaa 5460
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa 5515

<210> 6

<211> 6131

<212> DNA

<213> Homo sapiens

<400> 6

gaattctagg cccagttctg tgtttcccct gtgtgttcct aggcaggtca gtttccctcc 60
atgggcctct gtaagatgag gagttggaga ggtacattct caggctactt tcaactccca 120
gccaagtgac tcaagagtcc caggcagcac cagcacccct atctccaagg cctcctgatg 180
tgtgtctcta tttagaactt aatccaacct acccaacatc agatcagtgt cttaccagcc 240
caaggtccct ggggagcctc ctagagggag agagccctgc ccacccagat tgagggtaaa 300
ggcctccccg tgctcatttt tgtaccacca cagtgcttgg cacatggtag acatcaaaat 360
gtgtgtgctg aaagtataat tgaagttgtg tatatatgtc agctagagtg tctggagggg 420
cagaaatgtg ggtctaaaac atacaaatgc tccaaatggg gtgtgggcaa gggtctgtct 480
acaccaggct gtgattacct gctcacatac atgtgtctat ctgagtaggg gtatgttatc 540
tatttttcta caccacaggg tgaggaacag gtatatgtgt gcatgtgtat gcatccgtgt 600

gtgtgtgtat gtgtgtgtgc atgagtgtgt gtgtgtgtgt ccaaagccac ctcttcaacc 660
tgtgccattt gtatctgtgt ctggcccaat gagagtgttg aaaggtgagc cacaagataa 720
aacagcaact tcctacctcc cttatcaaga cagctgtctg acctacctcc ccttggccac 780
tcttgggatt actggggttg gcttcagtat tttcagattt ttcagaaggg gaggagaatg 840
cttgagtctc atccaggaac ttaggcagtt ctcagcactg cctgctcctc ctccctcaaa 900
taaccaagtc tgaagaccag gagagaaagc cgctggtgga ctggtcacct gtctggcagt 960
gggaggagga gagtgagagg tttctaggta ggaatccaga cttagaccct cccctccacc 1020
cccagatggg tggtgcacag gctcatctcg cggcccctcc ccactccacc ctaacatgga 1080
tacgccccca acaaccaagg aaagatctcc catcggctga ctccacagat acacacatgt 1140
ccccacagac acacacacgc ccatgcagag gcacagacat ccaggcacat ctttcccttt 1200
ctctgtcttt cccttggttt gaatttcgtt tagccacata tgttgtgtgt gcgtgagggt 1260
gggtggggga ggggcagaca gggatgaggg atggcatggt gccaacatct acctatgggg 1320
ctcgggccag ggacgcccct tacagccatc ctgggagggg gtctcagctg tccctttgtg 1380
gccaagggga ccctcctggg gagtgggggc aagcacagag gtcctttctc cccaacccgg 1440
ggtctggtcc ctgacccacc ttgggggcct gcaggggagg aaatggacag agcgggaccc 1500
tgagggagca tagaattggc caccacgagc ccccagtgtc cagccttgcc accccattgt 1560
tcccgtgagg gggtctctat atacaggggg caactcctcc caccttcctc tcaatccctg 1620
ctttccctgc gttgggcggg gaggggaggg cggcagaaat atttatttat ttcctttatt 1680
tatttaattt tttttttttt tttttggagt agagagtgac agatggcggc gggtcccggg 1740
ggagccggct ctcccccagt gcagacgcat gccaatcacc gtctctcatg tgatagctgc 1800
tgcccgtgac gtgccaagcc catatggcct ggcatagagg ctggtacccc gcctggtaga 1860
gatgccacac tcgctccgcg gttcgcatgg cgctctgaag acgccggcgc ccgccgcctt 1920
gaggagccgc tgccccccgct ccctgaagat gggggaacaa tgaaataagc gagaagatcc 1980
ctcttctccc ccctctctct cttgcccccct cccccccctcc cctcccctct ccccttgact 2040
cctctccgag gtaagttgtc cgaaagggag cgagatctga cccgccggtt gggaggaggg 2100
gcggcagctt cggccgacag gagggtcctc aaatacctcc ttcctgggat gatgcccccc 2160
tcattgggtg ggcatcggag gggccccagg ttctctctcc cttaggggct gcagcccagg 2220
gggctgcaga ggaggtgtct ctgcctgcga tgggctcggt gggggggaa ggcaggatca 2280
cggaggggga tatgcgaaga ggccgagacg gaggacccct ccatggttgt cccaaaaagc 2340
ctgccacctt tccccaccac cgaaaaaagg gaagcaaaca aacaaatttg gatttttccc 2400
ccatcaatcc caaaatacaa cgagatctga agagccttgt gggagggagt cagcttgaag 2460
ggggaagggg gtccctgacc gcagagggga cggactgggc tcgcttctct cagtctcctc 2520
cccacgcccc gctgcttcag tcctcgccgc ccagagccgg ctccgggagc tggggacgca 2580
tcggctagag gagacgatcc tcccgcctct ggaattgggg gtgcgggggt gggggccgag 2640
caaggggcgg cgcgcagcca agttgcaaat tggattaggg agcgtggggg tgagagccac 2700
gggaggggtg agggagctgg gccgggggc ccgggccgcg agagcgcgga gcggggcagc 2760
tgtccccacc ggcggccgac cagcctctct ccaccgccag gagagaacgg gctttcaggg 2820
cgagcgcgcc gcctcccctg gcaaagatat ctggtcccta aaacccccac ccggtccctg 2880
ccctgaccct gagaagaagc aggcgcgggg agcagccccc cattcaagcg aggggcggag 2940
ccgggggccca gcgccgggga gagggcctgg gccgagatcc caggccggca gccgggtagg 3000
gctgggccgg ctctgggcgg ggcaggcggc ggaggtgggc atccagggta gcctaggcag 3060
gagcccgcac gagactcggg ggtggaggag ggttgtgggg gggcgtcggt accccagcgc 3120
gcccctcact ttgtgctgtc tgtctcccct tcccgcccgc ggggcgccct caggcaccat 3180
gctgacccgc ctgttcagcg agcccggcct tctctcggac gtgcccaagt tcgccagctg 3240
gggcgacggc gaagacgacg agccgaggag cgacaagggc gacgcgccgc caccgccacc 3300
gcctgcgccc gggccagggg ctccggggcc agcccgggcg gccaagccag tccctctccg 3360
tggagaagag gggacggagg ccacgttggc cgaggtcaag gaggaaggcg agctgggggg 3420
agaggaggag gaggaagagg aggaggaaga aggactggac gaggcggagg gcgagcggcc 3480
caagaagcgc gggcccaaga agcgcaagat gaccaaggcg cgcttggagc gctccaagct 3540
tcggcggcag aaggcgaacg cgcgggagcg caaccgcatg cacgacctga acgcagccct 3600
ggacaacctg cgcaaggtgg tgccctgcta ctccaagacg cagaagctgt ccaagatcga 3660
gacgctgcgc ctagccaaga actatatctg ggcgctctcg gagatcctgc gctccggcaa 3720
gcggccagac ctagtgtcct acgtgcagac tctgtgcaag ggtctgtcgc agcccaccac 3780
caatctggtg gccggctgtc tgcagctcaa ctctcgcaac ttcctcacgg agcaaggcgc 3840
cgacggtgcc ggccgcttcc acggctcggg cggcccgttc gccatgcacc cctacccgta 3900
cccgtgctcg cgcctggcgg gcgcacagtg ccaggcggcc ggcggcctgg gcggcggcgc 3960
ggcgcacgcc ctgcggaccc acggctactg cgccgcctac gagacgctgt atgcggcggc 4020
aggcggtggc ggcgcgagcc cggactacaa cagctccgag tacgagggcc cgctcagccc 4080
cccgctctgt ctcaatggca acttctcact caagcaggac tcctcgcccg accacgagaa 4140
aagctaccac tactctatgc actactcggc gctgcccggt tcgcgccacg gccacgggct 4200
agtcttcggc tcgtcggctg tgcgcggggg cgtccactcg gagaatctct tgtcttacga 4260
tatgcacctt caccacgacc ggggccccat gtacgaggag ctcaatgcgt tttttcataa 4320
ctgagacttc gcgccggctc ccttcttttt cttttgcctt tgcccgcccc cctgtcccca 4380
gcccccagca gcgcagggta cacccccatc ctaccccggc gccgggcgcg gggagcgggc 4440
caccggtcct gccgctctcc tggggcagcg cagtcctgtt acctgtgggt ggcctgtccc 4500
aggggcctcg cttcccccag gggactcgcc ttctctctcc ccaaggggtt ccctcctcct 4560

```
ctctcccaag gagtgcttct ccagggacct ctctccgggg gctccctgga ggcacccctc   4620
ccccattccc aatatcttcg ctgaggtttc ctcctccccc tcctccctgc aggcccaagg   4680
cgttggtaag ggggcagctg agcaatggaa cgcgtttccc cctctcatta ttattttaaa   4740
aacagacacc cagctgccga ggcaaaaagg agccaggcgc tccctctttc ttgaagaggg   4800
tagtattttg ggcgccggag cccgggcctg gaacgccctc acccgcaacc tccagtctcc   4860
gcgttttgcg attttaattt tggcgggagg ggaagtggat tgagaggaaa gagagaggcc   4920
aagacaattt gtaactagaa tccgtttttc ccttttcctt tttttaaaca aacaaacata   4980
caaaaaaaaa aaaaaaaaaa aaaaaaaaaa aagctaagag gcgacggaag ccgaacgcag   5040
agtccggatc ggagagaaaa cgcagtaagg acttttagaa gcaataaaag gcaaaaaaaa   5100
caaaaaacaa aaaaacaaac aaaaaaaaac cactactacc aataatcaaa gacacaaata   5160
tctatgcaag gaggctccac tgagcctcgc ggcccggccc ggccccggga tgccccgccc   5220
ggcctgcggg ccgccccgcc cgagcgcgga tctgtgcact ttggtgaagt gggggcccgc   5280
gccgcccccct cccccctcccc aggttcttac aatcagtgac tcggagattt ggggccccag   5340
tgccactgcc ctcccccgcc ccgtccccgt tgtgcgtcat gctgtttttt aaaaacctgt   5400
ttccaaattt gtatggaatg gcaaactgtt ggggggtcgg tttggggagg gagggtttgc   5460
atgaaagaca cacgcacacc acaccgcacg cacaagcagg cccggcgccg gcgtccgggg   5520
ggcagaagga ggtgagctcg ccggctcctc ctccccgcgg ccattctgtc ccctcctggg   5580
gtgaggggtg gggatggaga cctgggggca gccccacccc tgcccggact gtgcctcggt   5640
gggtgccacc tggcgatttc cggtgtctgg agagagtatt ttttggtcca aggagtcctc   5700
ttggctttag ctggtgggtg ggcggggaga ggtctgaggg ctcctactgg aggttccccc   5760
aaaaaggggc aaaaggagac cctctgccca ccggaggcag gggatcaggc atccaaatac   5820
acgatgcaaa aatgcaatcc cacaggcgac acacccacac actcacccac acacacgcaa   5880
ttttaccttc ctcttgtagc gaagatgaaa ctcccgtcgg acacccgaag tgcattgcgt   5940
gtttctgttc agtttaatga cgattaataa atatttatgt aaatgagatg caaagccgga   6000
ccggtttctc acggtggcct catttcattg agggggggaga gaaggtttga gctggggctg   6060
gggtgatgaa ggcagagtgt caagtgactg tgcagaggcc aaacagaggg acttcccagc   6120
aaaaagcact g                                                        6131
```

<210> 7

<211> 2020

<212> DNA

<213> Homo sapiens

<400> 7

```
gctactgagg ccgcggagcc ggactgcggt tggggcggga agagccgggg ccgtggctga     60
catggagcag ccctgctgct gaggccgcgc cctccccgcc ctgaggtggg ggcccaccag    120
gatgagcaag ctgcccaggg agctgacccg agacttggag cgcagcctgc ctgccgtggc    180
ctccctgggc tcctcactgt cccacagcca gagcctctcc tcgcacctcc ttccgccgcc    240
tgagaagcga agggccatct ctgatgtccg ccgcaccttc tgtctcttcg tcaccttcga    300
cctgctcttc atctccctgc tctggatcat cgaactgaat accaacacag gcatccgtaa    360
gaacttggag caggagatca tccagtacaa ctttaaaact tccttcttcg acatctttgt    420
cctggccttc ttccgcttct ctggactgct cctaggctat gccgtgctgc agctccggca    480
ctggtgggtg attgcggtca cgacgctggt gtccagtgca ttcctcattg tcaaggtcat    540
cctctctgag ctgctcagca aaggggcatt tggctacctg ctccccatcg tctcttttgt    600
cctcgcctgg ttggagacct ggttccttga cttcaaagtc ctaccccagg aagctgaaga    660
ggagcgatgg tatcttgccg cccaggttgc tgttgcccgt ggacccctgc tgttctccgg    720
tgctctgtcc gagggacagt tctattcacc cccagaatcc tttgcagggt ctgacaatga    780
atcagatgaa gaagttgctg ggaagaaaag tttctctgct caggagcggg agtacatccg    840
ccaggggaag gaggccacgg cagtggtgga ccagatcttg gcccaggaag agaactggaa    900
gtttgagaag aataatgaat atggggacac cgtgtacacc attgaagttc cctttcacgg    960
caagacgttt atcctgaaga ccttcctgcc ctgtcctgcg gagctcgtgt accaggaggt   1020
gatcctgcag cccgagagga tggtgctgtg gaacaagaca gtgactgcct gccagatcct   1080
gcagcgagtg gaagacaaca ccctcatctc ctatgacgtg tctgcagggg ctgcgggcgg   1140
cgtggtctcc ccaagggact tcgtgaatgt ccggcgcatt gagcggcgca gggaccgata   1200
cttgtcatca gggatcgcca cctcacacag tgccaagccc ccgacgcaca aatatgtccg   1260
gggagagaat ggccctgggg gcttcatcgt gctcaagtcg gccagtaacc cccgtgtttg   1320
cacctttgtc tggattctta atacagatct caagggccgc ctgccccggt acctcatcca   1380
ccagagcctc gcggccacca tgtttgaatt tgcctttcac ctgcgacagc gcatcagcga   1440
gctgggggcc cgggcgtgac tgtgccccct cccaccctgc gggccagggt cctgtcgcca   1500
ccacttccag agccagaaag ggtgccagtt gggctcgcac tgcccacatg ggacctggcc   1560
ccaggctgtc accctccacc gagccacgca gtgcctggag ttgactgact gagcaggctg   1620
tggggtggag cactggactc cggggcccca ctggctggag gaagtggggt ctggcctgtt   1680
```

gatgtttaca tggcgccctg cctcctggag gaccagattg ctctgcccca ccttgccagg 1740
gcagggtctg ggctgggcac ctgacttggc tggggaggac cagggccctg ggcagggcag 1800
ggcagcctgt cacccgtgtg aagatgaagg ggctcttcat ctgcctgcgc tctcgtcggt 1860
tttttttagga ttattgaaag agtctgggac ccttgttggg gagtgggtgg caggtggggg 1920
tgggctgctg gccatgaatc tctgcctctc ccaggctgtc cccctcctcc cagggcctcc 1980
tgggggacct ttgtattaag ccaattaaaa acatgaattt 2020

<210> 8

<211> 1730

<212> DNA

<213> Homo sapiens

<400> 8

gtggtgaggg tgactgggga ctaggcacta ggcctttggt gcaggcgcct gaggacktgg 60
ttgcactctc ccttctgggg atatgccctt gagcccaggc agaggagagc acagcccagg 120
gcaggacctg gcagccctgg tacagagccc agagggggca tcagttcctg ctggtcctgc 180
tctgtttaca gacaasctgc tgtcctccct gcaaagggga gtgggtgggg cagagggcaa 240
ktgccagggg ggcacaaggc tgggcatgtg gctggcatga gacggtgtct gagtaatgtc 300
aggcacctgg aggcattgac cccaggacct tggaccccag acctctgacc gtggggcagc 360
cagcgtccag gtaccccaac ccctgccctg ggtccggcgt ccccccatta gtgagtcttg 420
gctctactta tagcatctga caccagaggg gccgaaaata gcccctggag aaggggggagg 480
agggggctat ttaaagggcc tgggaggggga gagagaatga ggagtgatca tggctacctc 540
agagctgagc tgcgaggtgt cggaggagaa ctgtgagcgc cgggaggcct tctgggcaga 600
atggaaggat ctgacactgt ccacacggcc cgaggagggg tgagtgtggg tctgctagag 660
tccctgcctc tgctccccca gagcaccctc actgagccat gaggccagag catgaagccc 720
tggagaaatt tctgggggtg ggggcaggaa gaatgcccca tggggagagc aaaggggaac 780
cacccttcct gcccccaggt cccagcagcc caggggagcc ccccacccag cctgtgccca 840
gagagcaaca gctcccagga gctcactgcc cctcccctct ccccagctgc tccctgcatg 900
aggaggacac ccagagacat gagacctacc accagcaggg gcagtgccag gtgctggtgc 960
agcgctcgcc ctggctgatg atgcggatgg gcatcctcgg ccgtgggctg caggagtacc 1020
agctgcccta ccagcgggta ctgccgctgc ccatcttcac ccctgccaag atgggcgcca 1080
ccaaggagga gcgtgaggac acccccatcc agcttcagga gctgctggcg ctggagacag 1140
ccctgggtgg ccagtgtgtg gaccgccagg aggtggctga gatcacaaag cagctgcccc 1200
ctgtggtgcc tgtcagcaag cccggtgcmc ttcgtcgctc cctgtcccgc tccatgtccc 1260
aggaagcaca gagaggctga gagggactgt gacttgggct ccgctgtgcc cgccccctgg 1320
gctgggccct tcctggctag gacctgtgga ggggcagctc gctggcccat ggctgctttg 1380
tagtttgccc agagttgggg gctaggggag gggggagcca gaggccagga tgcctgagcc 1440
ccctgagttc ccaaagggag ggtggcagag acagtgggca ctaagggtgg agagttgggg 1500
gccagcacag ctgaggaccc tcagccccag gagaagggac aaaaggtact ggtgagggca 1560
agaggtgcct gggaggagtg gccctgatcc aggaaaatgt gaggggaatc tggaacgctc 1620
taggcagaag aagctgggag ggagggggag gtgaaaaggg cagaggcaag gatggtgggg 1680
cccccagcac cctctgttag tgccgcaata aatgctcaat catgtgccag 1730

<210> 9

<211> 3799

<212> DNA

<213> Homo sapiens

<400> 9

ctggcactgg gtggtaacca gcaagccagc tggcatccgc atccagggtt tgtttcaatg 60
atgtctcgtg gagaatatgg aggggctggt gccaggactg tccttggctt tgcctcgggg 120
tgtgaacggg gtcagtgacc tctaaaacta acctgcctct cagttctgaa tccagacaga 180
atcaatcctc agctgtgtct cgctccacac cccctgccct ggaagccagg gaaggttgga 240
ggtgctaggg ggtcaggctc ccctctgtga cccctgcagc tgttgtggtg actcatgtcc 300
caacctagct gcctctccca aggagacttt cccctgggac aagggggagg gaatggcatg 360
gaggaggccc acatcaagcg gggccaggaa cccacggtgg caggagctgg gctggtgacc 420
tacccagggc agaagggccc gggactcatc cagaggggaa ggaaggggtc ttcaggaaga 480
ccacggagat gccacaggca gaattggctt cccatctggg agataggtgg ggagaccctg 540

gcattttgac agccagaacc tggggtgctg agcagaatct tcatgcctgg cctggccgcc 600
ttcggaggga agctggaggg ttgggtgcga gaggagtggg gtcagagccc ctacatccgc 660
aggaccccaa atcggctggg ccccaaggcc cggactgcgc tccccggtgg ccccggcggc 720
cctccgcgaa tgcgtcctgc ccctcccctg cccaagccct ctgccctcac ccgggtccgg 780
cgccgccccc gaagtggcgg gaacaacccg aacccgaacc ttctgtcctc gggagccccc 840
agataagcgg ctgggaaccc gcgggcccg cagggaggc ccggctgttc cgcccgctaa 900
gtgcattagc acagctcacc tcccctatcg cgcctgccat cggacgggca gtgccgcgcc 960
ctgctctggg gcccccggag cgaccacagc ggaggccgga acggactgtc ctttctgggg 1020
cggggtgggg agggggtgtc gctggagggc ccggtggcat agcaacggac gagagaggcc 1080
tggaggaggg gcggggaggg ggagttgtgt ggcagttcta agggaagggt gggtgctggg 1140
acgggtgtcc gggagggagg ggagcctggc ggggtctggg gcctcgtcgc ggagggcgct 1200
gcgaggggga aactggggaa agggcctaat tccccagtct ccacctcgaa tcaggaaaga 1260
gaaggggcgg gctgctgggc aaaagaggtg aatggctgcg gggggctgga gaagagagat 1320
gggaggggcc ggccggcggg ggtgaggggg tctaaagatt gtgggggtga ggaactgagg 1380
gtgggggcg cccagaggcg ggactcgggg cggggcaggc gaggcggagg gcgagggctg 1440
cgggagcaag tacggagccg gggtgtgggg ggacgattgc cgctgcagcc gccgccccac 1500
tcacctccgg tgtgtctgca gcccggacac taagggagat ggatgaatgg gtggggagga 1560
tgcggcgcac atggccccgg gcggctcggc ggtcagctgc cgcccccaca gcggaccggt 1620
cggggcgggg gtcgggcggt agaaaaaagg gccgcgaggc gagcggggca ctgggcggac 1680
cgcggcggca gcatgagcgg cgcagaccgt agccccaatg cgggcgcagc ccctgactcg 1740
gccccgggcc aggcggcggt ggcttcggcc taccagcgct tcgagccgcg cgcctacctc 1800
cgcaacaact acgcgccccc tcgcggggac ctgtgcaacc cgaacggcgt cgggccgtgg 1860
aagctgcgct gcttggcgca gaccttcgcc accggtgagc gggggaaact gaggcacgag 1920
ggacaagagg tcgtcgggga gtgaaagcag gcgcagggaa ataaaaagaa ggaaagggag 1980
acagaccagg cgcctaacag atggggacca agaaacaaga gatagctgag aggtgcaaac 2040
agaagagaaa aaggagcaac atcccttagg agaggggcag aggagagaga ggtggagaga 2100
gggggcggag agtgctcaga attgagagct aaggtggggg atgcaggaca gactgaggtg 2160
gagatgcata ggaggaaatg gaggcagatg tgggacaggg gtgagaaact ccaggatttc 2220
ctcgctgagc ctggctggta ggtatagttg ttttctttct ttttctttat tttattttca 2280
tttatttact tatttttatt ttttatttgt tttgagacgg agtttcgctc ttgttgccca 2340
ggctggagta caatggcgcc atctcggctc actgcaacct ccgcctcccc gggttcaagc 2400
gattctcttg cctcagcttc cctagtagct gggattacag gcatgcgccc ccatgcctgg 2460
ctaatttatt tgtattttta gtagagacgg gacttctcca tgttggtcag gctggtctcg 2520
aactcccaac cttaggatcc acccaccccg gcctcccaaa gtgctgggat tacaggtgtg 2580
agccactgcg cccggccagt aggtatagtc ttctagatgt gaaacctgag tctcagagcg 2640
gtgaagttcc cttccgaagg gcagcccatg ttggagctgg gttcagtcta actctggggc 2700
caatgctttt tccagatgga gacacatttg cagaggagaa ggaagaacta gagagaggca 2760
gggagatgca ggggagggaa gggtaaggag gcaggggctg cctgggctgg ctggcaccag 2820
gaccctcttc ctctgccctg cccaggtgaa gtgtccggac gcaccctcat cgacattggt 2880
tcaggcccca ccgtgtacca gctgctcagt gcctgcagcc actttgagga catcaccatg 2940
acagatttcc tggaggtcaa ccgccaggag ctgggggcgct ggctgcagga ggagccgggg 3000
gccttcaact ggagcatgta cagccaacat gcctgcctca ttgagggcaa ggggtaagga 3060
ctgggggtg agggttgggg aggaggcttc ccatagagtg gctggttggg gcaacagagg 3120
cctgagcgta gaacagcctt gagccctgcc ttgtgcctcc tgcacaggga atgctggcag 3180
gataaggagc gccagctgcg agccagggtg aaacggtcc tgcccatcga cgtgcaccag 3240
ccccagcccc tgggtgctgg gagcccagct cccctgcctg ctgacgccct ggtctctgcc 3300
ttctgcttgg aggctgtgag cccagatctt gccagctttc agcgggccct ggaccacatc 3360
accacgctgc tgaggcctgg ggggcacctc ctcctcatcg ggggccctgga ggagtcgtgg 3420
tacctggctg ggaggccag gctgacggtg gtgccagtgt ctgaggagga ggtgagggag 3480
gccctggtgc gtagtggcta caaggtccgg gacctccgca cctatatcat gcctgcccac 3540
cttcagacag gcgtagatga tgtcaagggc gtcttcttcg cctgggctca gaaggttggg 3600
ctgtgagggc tgtacctggt gccctgtggc ccccacccac ctggattccc tgttctttga 3660
agtggcacct aataaagaaa taataccctg ccgctgcggt cagtgctgtg tgtggctctc 3720
ctgggaagca gcaagggccc agagatctga gtgtccgggt aggggagaca ttcaccctag 3780
gcttttttc cagaagctt 3799

<210> 10

<211> 4530

<212> DNA

<213> Homo sapiens

```
<400>  10
aattctcgag ctcgtcgacc ggtcgacgag ctcgagggtc gacgagctcg agggcgcgcg      60
cccggccccc acccctcgca gcaccccgcg ccccgcgccc tcccagccgg gtccagccgg     120
agccatgggg ccggagccgc agtgagcacc atggagctgg cggccttgtg ccgctggggg     180
ctcctcctcg ccctcttgcc ccccggagcc gcgagcaccc aagtgtgcac cggcacagac     240
atgaagctgc ggctccctgc cagtcccgag acccacctgg acatgctccg ccacctctac     300
cagggctgcc aggtggtgca gggaaacctg gaactcacct acctgcccac caatgccagc     360
ctgtccttcc tgcaggatat ccaggaggtg cagggctacg tgctcatcgc tcacaaccaa     420
gtgaggcagg tcccactgca gaggctgcgg attgtgcgag gcacccagct ctttgaggac     480
aactatgccc tggccgtgct agacaatgga gacccgctga acaataccac ccctgtcaca     540
ggggcctccc caggaggcct gcgggagctg cagcttcgaa gcctcacaga gatcttgaaa     600
ggaggggtct tgatccagcg gaacccccag ctctgctacc aggacacgat tttgtggaag     660
gacatcttcc acaagaacaa ccagctggct ctcacactga tagacaccaa ccgctctcgg     720
gcctgccacc cctgttctcc gatgtgtaag ggctcccgct gctggggaga gagttctgag     780
gattgtcaga gcctgacgcg cactgtctgt gccggtggct gtgcccgctg caaggggcca     840
ctgcccactg actgctgcca tgagcagtgt gctgccggct gcacgggccc caagcactct     900
gactgcctgg cctgcctcca cttcaaccac agtggcatct gtgagctgca ctgcccagcc     960
ctggtcacct acaacacaga cacgtttgag tccatgccca atcccgaggg ccggtataca    1020
ttcggcgcca gctgtgtgac tgcctgtccc tacaactacc tttctacgga cgtgggatcc    1080
tgcaccctcg tctgccccct gcacaaccaa gaggtgacag cagaggatgg aacacagcgg    1140
tgtgagaagt gcagcaagcc ctgtgcccga gtgtgctatg gtctgggcat ggagcacttg    1200
cgagaggtga gggcagttac cagtgccaat atccaggagt ttgctggctg caagaagatc    1260
tttgggagcc tggcatttct gccggagagc tttgatgggg acccagcctc caacactgcc    1320
ccgctccagc cagagcagct ccaagtgttt gagactctgg aagagatcac aggttaccta    1380
tacatctcag catggccgga cagcctgcct gacctcagcg tcttccagaa cctgcaagta    1440
atccggggac gaattctgca caatggcgcc tactcgctga ccctgcaagg gctgggcatc    1500
agctggctgg ggctgcgctc actgagggaa ctgggcagtg gactggccct catccaccat    1560
aacacccacc tctgcttcgt gcacacggtg ccctgggacc agctctttcg gaacccgcac    1620
caagctctgc tccacactgc caaccggcca gaggacgagt gtgtgggcga gggcctggcc    1680
tgccaccagc tgtgcgcccg agggcactgc tggggtccag ggcccaccca gtgtgtcaac    1740
tgcagccagt tccttcgggg ccaggagtgc gtggaggaat gccgagtact gcaggggctc    1800
cccagggagt atgtgaatgc caggcactgt ttgccgtgcc accctgagtg tcagccccag    1860
aatggctcag tgacctgttt tggaccggag gctgaccagt gtgtggcctg tgcccactat    1920
aaggaccctc ccttctgcgt ggcccgctgc cccagcggtg tgaaacctga cctctcctac    1980
atgcccatct ggaagtttcc agatgaggag ggcgcatgcc agccttgccc catcaactgc    2040
acccactcct gtgtggacct ggatgacaag ggctgccccg ccgagcagag agccagccct    2100
ctgacgtcca tcgtctctgc ggtggttggc attctgctgg tcgtggtctt gggggtggtc    2160
tttgggatcc tcatcaagcg acggcagcag aagatccgga agtacacgat gcggagactg    2220
ctgcaggaaa cggagctggt ggagccgctg acacctagcg gagcgatgcc caaccaggcg    2280
cagatgcgga tcctgaaaga gacggagctg aggaaggtga aggtgcttgg atctggcgct    2340
tttggcacag tctacaaggg catctggatc cctgatgggg agaatgtgaa aattccagtg    2400
gccatcaaag tgttgaggga aaacacatcc cccaaagcca acaaagaaat cttagacgaa    2460
gcatacgtga tggctggtgt gggctcccca tatgtctccc gccttctggg catctgcctg    2520
acatccacgg tgcagctggt gacacagctt atgccctatg gctgcctctt agaccatgtc    2580
cgggaaaacc gcggacgcct gggctcccag gacctgctga actggtgtat gcagattgcc    2640
aaggggatga gctacctgga ggatgtgcgg ctcgtacaca gggacttggc cgctcggaac    2700
gtgctggtca agagtcccaa ccatgtcaaa attacagact tcgggctggc tcggctgctg    2760
gacattgacg agacagagta ccatgcagat gggggcaagg tgcccatcaa gtggatggcg    2820
ctggagtcca ttctccgccg gcggttcacc caccagagtg atgtgtggag ttatggtgtg    2880
actgtgtggg agctgatgac ttttggggcc aaaccttacg atgggatccc agcccgggag    2940
atccctgacc tgctggaaaa gggggagcgg ctgccccagc cccccatctg caccattgat    3000
gtctacatga tcatggtcaa atgttggatg attgactctg aatgtcggcc aagattccgg    3060
gagttggtgt ctgaattctc ccgcatggcc agggaccccc agcgctttgt ggtcatccag    3120
aatgaggact tgggcccagc cagtcccttg gacagcacct tctaccgctc actgctggag    3180
gacgatgaca tgggggacct ggtggatgct gaggagtatc tggtacccca gcagggcttc    3240
ttctgtccag accctgcccc gggcgctggg ggcatggtcc accacaggca ccgcagctca    3300
tctaccagga gtggcggtgg ggacctgaca ctagggctgg agccctctga agaggaggcc    3360
cccaggtctc cactggcacc ctccgaaggg gctggctccg atgtatttga tggtgacctg    3420
ggaatggggg cagccaaggg gctgcaaagc ctccccacac atgaccccag ccctctacag    3480
cggtacagtg aggaccccac agtacccctg ccctctgaga ctgatggcta cgttgccccc    3540
ctgacctgca gcccccagcc tgaatatgtg aaccagccag atgttcggcc ccagccccct    3600
tcgccccgag agggccctct gcctgctgcc cgacctgctg gtgccactct ggaaagggcc    3660
aagactctct ccccagggaa gaatggggtc gtcaaagacg tttttgcctt tgggggtgcc    3720
gtggagaacc ccgagtactt gacaccccag ggaggagctg cccctcagcc ccaccctcct    3780
cctgccttca gcccagcctt cgacaacctc tattactggg accaggaccc accagagcgg    3840
ggggctccac ccagcacctt caaagggaca cctacggcag agaacccaga gtacctgggt    3900
```

ctggacgtgc cagtgtgaac cagaaggcca agtccgcaga agccctgatg tgtcctcagg 3960
gagcagggaa ggcctgactt ctgctggcat caagaggtgg gagggccctc cgaccacttc 4020
caggggaacc tgccatgcca ggaacctgtc ctaaggaacc ttccttcctg cttgagttcc 4080
cagatggctg gaaggggtcc agcctcgttg gaagaggaac agcactgggg agtctttgtg 4140
gattctgagg ccctgcccaa tgagactcta gggtccagtg gatgccacag cccagcttgg 4200
ccctttcctt ccagatcctg ggtactgaaa gccttaggga agctggcctg agaggggaag 4260
cggccctaag ggagtgtcta agaacaaaag cgacccattc agagactgtc cctgaaacct 4320
agtactgccc cccatgagga aggaacagca atggtgtcag tatccaggct ttgtacagag 4380
tgcttttctg tttagttttt actttttttg ttttgttttt ttaaagacga aataaagacc 4440
caggggagaa tgggtgttgt atggggaggc aagtgtgggg ggtccttctc cacacccact 4500
ttgtccattt gcaaatatat tttggaaaac 4530

<210> 11

<211> 2205

<212> DNA

<213> Homo sapiens

<400> 11
cacagggctc ccccccgcct ctgacttctc tgtccgaagt cgggacaccc tcctaccacc 60
tgtagagaag cgggagtgga tctgaaataa aatccaggaa tctggggggtt cctagacgga 120
gccagacttc ggaacgggtg tcctgctact cctgctgggg ctcctccagg acaagggcac 180
acaactggtt ccgttaagcc cctctctcgc tcagacgcca tggagctgga tctgtctcca 240
cctcatctta gcagctctcc ggaagacctt tggccagccc ctgggacccc tcctgggact 300
ccccggcccc ctgatacccc tctgcctgag gaggtaaaga ggtcccagcc tctcctcatc 360
ccaaccaccg gcaggaaact tcgagaggag gagaggcgtg ccacctccct cccctctatc 420
cccaacccct tccctgagct ctgcagtcct ccctcacaga gcccaattct cgggggcccc 480
tccagtgcaa gggggctgct cccccgcgat gccagccgcc cccatgtagt aaaggtgtac 540
agtgaggatg gggcctgcag gtctgtggag gtggcagcag gtgccacagc tcgccacgtg 600
tgtgaaatgc tggtgcagcg agctcacgcc ttgagcgacg agacctgggg gctggtggag 660
tgccaccccc acctagcact ggagcggggt ttggaggacc acgagtccgt ggtggaagtg 720
caggctgcct ggcccgtggg cggagatagc cgcttcgtct tccggaaaaa cttcgccaag 780
tacgaactgt tcaagagctc ccccacactcc ctgttcccag aaaaaatggt ctccagctgt 840
ctcgatgcac acactggtat atcccatgaa gacctcatcc agaacttcct gaatgctggc 900
agctttcctg agatccaggg ctttctgcag ctgcggggtt caggacggaa gctttggaaa 960
cgctttttct gtttcttgcg ccgatctggc ctctattact ccaccaaggg cacctctaag 1020
gatccgaggc acctgcagta cgtggcagat gtgaacgagt ccaacgtgta cgtggtgacg 1080
cagggccgca agctctacgg gatgcccact gacttcggtt tctgtgtcaa gcccaacaag 1140
cttcgaaatg gacacaaggg gcttcggatc ttctgcagtg aagatgagca gagccgcacc 1200
tgctggctgg ctgccttccg cctcttcaag tacggggtgc agctgtacaa gaattaccag 1260
caggcacagt ctcgccatct gcatccatct tgtttgggct ccccaccctt gagaagtgcc 1320
tcagataata ccctggtggc catggacttc tctggccatg ctgggcgtgt cattgagaac 1380
ccccgggagg ctctgagtgt ggccctggag gaggcccagg cctggaggaa gaagacaaac 1440
caccgcctca gcctgcccat gccagcctcc ggcacgagcc tcagtgcagc catccaccgc 1500
acccaactct ggttccacgg gcgcatttcc cgtgaggaga gccagcggct tattggacag 1560
cagggcttgg tagacggcct gttcctggtc cgggagagtc agcggaaccc ccagggcttt 1620
gtcctctctt tgtgccacct gcagaaagtg aagcattatc tcatcctgcc gagcgaggag 1680
gagggtcgcc tgtacttcag catggatgat ggccagaccc gcttcactga cctgctgcag 1740
ctcgtggagt tccaccagct gaaccgcggc atcctgccgt gcttgctgcg ccattgctgc 1800
acgcgggtgg ccctctgacc aggccgtgga ctggctcatg cctcagcccg ccttcaggct 1860
gcccgccgcc cctccaccca tccagtggac tctggggcgc ggccacaggg gacgggatga 1920
ggagcgggag ggttccgcca ctccagtttt ctcctctgct tctttgcctc cctcagatag 1980
aaaacagccc ccactccagt ccactcctga cccctctcct caagggaagg ccttgggtgg 2040
ccccctctcc ttctcctagc tctggaggtg ctgctctagg gcagggaatt atgggagaag 2100
tgggggcagc ccaggcggtt tcacgcccca cactttgtac agaccgagag gccagttgat 2160
ctgctctgtt ttatactagt gacaataaag attatttttt gatac 2205

<210> 12

<211> 2177

<212> DNA

<213> Homo sapiens

<400> 12
gaattcgcgg ccgctggttt gcagctgctc cgtcatcgtg cggcccgacg ctatctcgcg 60
ctcgtgtgca ggcccggctc ggctcctggt ccccggtgcg agggttaacg cgaggccccg 120
gcctcggtcc ccggactagg ccgtgacccc gggtgccatg aagcaggagg gctcggcgcg 180
gcgccgcggc gcggacaagg cgaaaccgcc gcccggcgga ggagaacaag aaccccccacc 240
gccgccggcc ccccaggatg tggagatgaa agaggaggca gcgacgggtg gcgggtcaac 300
gggggaggca gacggcaaga cggcggcggc agcggttgag cactcccagc gagagctgga 360
cacagtcacc ttggaggaca tcaaggagca cgtgaaacag ctagagaaag cggtttcagg 420
caaggagccg agattcgtgc tgcgggccct gcggatgctg ccttccacat cacgccgcct 480
caaccactat gttctgtata aggctgtgca gggcttcttc acttcaaata atgccactcg 540
agactttttg ctccccttcc tggaagagcc catggacaca gaggctgatt tacagttccg 600
tccccgcacg ggaaaagctg cgtcgacacc cctcctgcct gaagtggaag cctatctcca 660
actcctcgtg gtcatcttca tgatgaacag caagcgctac aaagaggcac agaagatctc 720
tgatgatctg atgcagaaga tcagtactca gaaccgccgg gccctagacc ttgtagccgc 780
aaagtgttac tattatcacg cccgggtcta tgagttcctg gacaagctgg atgtggtgcg 840
cagcttcttg catgctcggc tccggacagc tacgcttcgg catgacgcag acgggcaggc 900
caccctgttg aacctcctgc tgcggaatta cctacactac agcttgtacg accaggctga 960
gaagctggtg tccaagtctg tgttcccaga gcaggccaac aacaatgagt gggccaggta 1020
cctctactac acagggcgaa tcaaagccat ccagctggag tactcagagg cccggagaac 1080
gatgaccaac gcccttcgca aggcccctca gcacacagct gtcggcttca aacagacggt 1140
gcacaagctt ctcatcgtgg tggagctgtt gctgggggag atccctgacc ggctgcagtt 1200
ccgccagccc tccctcaagc gctcactcat gccctatttc cttctgactc aagctgtcag 1260
gacaggaaac ctagccaagt tcaaccaggt cctggatcag tttggggaga agtttcaagc 1320
agatgggacc tacaccctaa ttatccggct gcggcacaac gtgattaaga caggtgtacg 1380
catgatcagc ctctcctatt cccgaatctc cttggctgac atcgcccaga agctgcagtt 1440
ggatagcccc gaagatgcag agttcattgt tgccaaggcc atccgggatg gtgtcattga 1500
ggccagcatc aaccacgaga agggctatgt ccaatccaag gagatgattg acatctattc 1560
cacccgagag ccccagctag ccttccacca gcgcatctcc ttctgcctag atatccacaa 1620
catgtctgtc aaggccatga ggtttcctcc caaatcgtac aacaaggact tggagtctgc 1680
agaggaacgg cgtgagcgag aacagcagga cttggagttt gccaaggaga tggcagaaga 1740
tgatgatgac agcttccctt gagctggggg gctggggagg ggtaggggga atggggacag 1800
gctctttccc ccttgggggt cccctgccca gggcactgtc cccatttcc cacacacagc 1860
tcatatgctg cattcgtgca gggggtgggg gtgctgggag ccagccaccc tgacctcccc 1920
cagggctcct ccccagccgg tgacttactg tacagcaggc aggagggtgg gcaggcaacc 1980
tccccgggca gggtcctggc cagcagtgtg ggagcaggag gggaaggata gttctgtgta 2040
ctcctttagg gagtggggga ctagaactgg gatgtcttgg cttgtatgtt ttttgaagct 2100
tcgattatga tttttaaaca ataaaaagtt ctcccaaaaa aaaaaaaaaa aaaaaaaaaa 2160
aaagcggccg cgaattc 2177
<210> 13

<211> 2960

<212> DNA

<213> Homo sapiens

<400> 13
ctgccgcttc caggcgtcta tcagcggctc agcctttgtt cagctgttct gttcaaacac 60
tctgggggcca ttcaggcctg ggtggggcag cgggaggaag ggagtttgag gggggcaagg 120
cgacgtcaaa ggaggatcag agattccaca atttcacaaa actttcgcaa acagcttttt 180
gttccaaccc ccctgcattg tcttggacac caaatttgca taaatcctgg gaagttatta 240
ctaagcctta gtcgtggccc caggtaattt cctcccaggc ctccatgggg ttatgtataa 300
agggcccccct agagctgggc cccaaaacag cccggagcct gcagcccagc cccacccaga 360
cccatggctg gacctgccac ccagagcccc atgaagctga tgggtgagtg tcttggccca 420
ggatgggaga gccgcctgcc ctggcatggg agggaggctg gtgtgacaga ggggctgggg 480
atccccgttc tgggaatggg gattaaaggc acccagtgtc cccgagaggg cctcaggtgg 540
tagggaacag catgtctcct gagcccgctc tgtccccagc cctgcagctg ctgctgtggc 600
acagtgcact ctggacagtg caggaagcca ccccccctggg ccctgccagc tccctgcccc 660
agagcttcct gctcaagtgc ttagagcaag tgaggaagat ccagggcgat ggcgcagcgc 720
tccaggagaa gctggtgagt gaggtgggtg agagggctgt ggagggaagc ccggtgggga 780
gagctaaggg ggatggaact gcagggccaa catcctctgg aagggacatg ggagaatatt 840

aggagcagtg gagctgggga aggctgggaa gggacttggg gaggaggacc ttggtgggga 900
cagtgctcgg gagggctggc tgggatggga gtggaggcat cacattcagg agaaagggca 960
agggcccctg tgagatcaga gagtgggggt gcagggcaga gaggaactga acagcctggc 1020
aggacatgga gggaggggaa agaccagaga gtcggggagg acccgggaag gagcggcgac 1080
ccggccacgg cgagtctcac tcagcatcct tccatcccca gtgtgccacc tacaagctgt 1140
gccaccccga ggagctggtg ctgctcggac actctctggg catcccctgg gctcccctga 1200
gcagctgccc cagccaggcc ctgcagctgg tgagtgtcag gaaaggataa ggctaatgag 1260
gagggggaag gagaggagga acacccatgg gctcccccat gtctccaggt tccaagctgg 1320
gggcctgacg tatctcaggc agcaccccct aactcttccg ctctgtctca caggcaggct 1380
gcttgagcca actccatagc ggcctttcc tctaccaggg gctcctgcag gccctggaag 1440
ggatctcccc cgagttgggt cccaccttgg acacactgca gctggacgtc gccgactttg 1500
ccaccaccat ctggcagcag gtgagccttg ttgggcaggg tggccaaggt cgtgctggca 1560
ttctgggcac cacagccggg cctgtgtatg ggccctgtcc atgctgtcag cccccagcat 1620
ttcctcattt gtaataacgc ccactcagaa gggcccaacc actgatcaca gctttccccc 1680
acagatggaa gaactgggaa tggcccctgc cctgcagccc acccagggtg ccatgccggc 1740
cttcgcctct gctttccagc gccgggcagg aggggtcctg gttgcctccc atctgcagag 1800
cttcctggag gtgtcgtacc gcgttctacg ccaccttgcc cagccctgag ccaagccctc 1860
cccatcccat gtatttatct ctatttaata tttatgtcta tttaagcctc atatttaaag 1920
acagggaaga gcagaacgga gccccaggcc tctgtgtcct tccctgcatt tctgagtttc 1980
attctcctgc ctgtagcagt gagaaaaagc tcctgtcctc ccatcccctg gactgggagg 2040
tagataggta aataccaagt atttattact atgactgctc cccagccctg gctctgcaat 2100
gggcactggg atgagccgct gtgagcccct ggtcctgagg gtccccacct gggacccttg 2160
agagtatcag gtctcccacg tgggagacaa gaaatccctg tttaatattt aaacagcagt 2220
gttccccatc tgggtccttg caccctcac tctggcctca gccgactgca cagcggcccc 2280
tgcatcccct tggctgtgag gcccctggac aagcagaggt ggccagagct gggaggcatg 2340
gccctggggt cccacgaatt tgctggggaa tctcgttttt cttcttaaga cttttgggac 2400
atggtttgac tcccgaacat caccgacgtg tctcctgttt ttctgggtgg cctcgggaca 2460
cctgccctgc ccccacgagg gtcaggactg tgactctttt tagggccagg caggtgcctg 2520
gacatttgcc ttgctggatg gggactgggg atgtgggagg gagcagacag gaggaatcat 2580
gtcaggcctg tgtgtgaaag gaagctccac tgtcaccctc cacctcttca ccccccactc 2640
accagtgtcc cctccactgt cacattgtaa ctgaacttca ggataataaa gtgtttgcct 2700
ccagtcacgt ccttcctcct tcttgagtcc agctggtgcc tggccagggg ctggggaggt 2760
ggctgaaggg tgggagaggc cagagggagg tcggggagga ggtctgggga ggaggtccag 2820
ggaggaggag gaaagttctc aagttcgtct gacattcatt ccgttagcac atatttatct 2880
gagcacctac tctgtgcaga cgctgggcta agtgctgggg acacagcagg gaacaaggca 2940
gacatggaat ctgcactcga 2960

<210> 14

<211> 850

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (3)..(4)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (9)..(9)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (11)..(11)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (18)..(18)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (202)..(202)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (205)..(205)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (273)..(273)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (327)..(327)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (367)..(367)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (581)..(581)

<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (599)..(599)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (628)..(628)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (673)..(673)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (675)..(675)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (682)..(682)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (693)..(693)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (698)..(698)
<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (700)..(700)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (720)..(720)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (730)..(730)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (734)..(734)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (742)..(743)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (746)..(746)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (748)..(748)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (752)..(752)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (762)..(762)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (767)..(767)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (777)..(777)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (783)..(784)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (789)..(789)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (794)..(794)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (797)..(798)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (803)..(805)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (810)..(810)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (817)..(817)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (826)..(827)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (831)..(832)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (834)..(834)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (837)..(838)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (840)..(840)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (844)..(844)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (846)..(848)

<223> n=a, c, g or t

<400> 14
```
ttnnctttnt ngccatgncc agttcaactc agcctctcag ttccacacgg acaacatgcg      60
ggaccctctg aaccgagtcc tggccaacct gttcctgctc atctcctcca tcctggggtc     120
tcgcaccgct ggcccccaca cccagttcgt gcagtggttc atggaggagt gtgtggactg     180
cctggagcag ggtggccgtg gnagngtcct gcagttcatg cccttcacca ccgtgtcgga     240
actggtgaag gtgtcagcca tgtctagccc canggtggtt ctggccatca cggacctcag     300
cctgcccctg ggccgccagg tggctgntaa agccattgct gcactctgag ggcttggca      360
tggccgnagt gggggctggg gactggcgca gccccaggcg cctccaaggg aagcagtgag     420
gaaagatgag gcatcgtgcc tcacatccgt tccacatggt gcaagagcct ctagcggctt     480
ccagttcccc gctcctgact cctgactcca ggatgtctcc cggtttcttc ttttcaaaat     540
tttcctctcc atcttgctgg caactgagga gagtgagcag nctggaccac aagcccagng     600
ggtcacccct gtgttgcgcc cgcccagncc aggagtagtc ttacctcttg aggaactttc     660
ttggatggaa agngngtttt tntgtgttgt gtntgtgnan gtgtttttcg gggtttttn      720
gggcaatatn ttangggaat cnnccntncg cncatttttt cnttagagct ccccggngga     780
aanntcttna tccnctnnct ttnnnctccn tcacctncct tctttnntct nntnttnncn     840
tccncnnncc                                                            850
```
<210> 15

<211> 2309

<212> DNA

<213> Homo sapiens

<400> 15
```
ccccgggcgc aggaggcggg cggcccggcc ccaccggccc cccatggacg cccccagcac      60
ggggcgctga gaccccccgcg tcgctgccca gcccggtccg gcgcgccacg ccagggatct    120
ctggacagga caagactccg aagctactcc cccagcacac agcccgggac ccacaaaccc     180
agcttgcccc cagccctccc acctgccact ccctggcccc tcccaccgcc cgcccccctt     240
ggggcgcagg gcatggtgtg aaaggccaag tgctgaggcg ggtatcatgg gtgctgtgcc     300
ctagggcctg ggtggcaggg ggtgggtggc ctgtgggtgt gccggggggg ccagtgtgcc     360
cacccccagtc tcttggcgtg ctggagggca tcctggatgg aattgaagtg aatggaacag    420
aagccaagca aggtggagtg tgggtcagac ccagaggaga acagtgccag gtcaccagat     480
ggaaagcgaa aaagaaagaa cggccaatgt tccctgaaaa ccagcatgtc agggtatatc     540
cctagttacc tggacaaaga cgagcagtgt gtcgtgtgtg gggacaaggc aactggttat     600
cactaccgct gtatcacttg tgagggctgc aagggcttct ttcgccgcac aatccagaag     660
```

```
aacctccatc ccacctattc ctgcaaatat gacagctgct gtgtcattga caagatcacc    720
cgcaatcagt gccagctgtg ccgcttcaag aagtgcatcg ccgtgggcat ggccatggac    780
ttggttctag atgactcgaa gcgggtggcc aagcgtaagc tgattgagca gaaccgggag    840
cggcggcgga aggaggagat gatccgatca ctgcagcagc gaccagagcc cactcctgaa    900
gagtgggatc tgatccacat tgccacagag gcccatcgca gcaccaatgc ccagggcagc    960
cattggaaac agaggcggaa attcctgccc gatgacattg gccagtcacc cattgtctcc   1020
atgccggacg gagacaaggt ggacctggaa gccttcagcg agtttaccaa gatcatcacc   1080
ccggccatca cccgtgtggt ggactttgcc aaaaaactgc ccatgttctc cgagctgcct   1140
tgcgaagacc agatcatcct cctgaagggg tgctgcatgg agatcatgtc cctgcgggcg   1200
gctgtccgct acgaccctga gagcgacacc ctgacgctga gtggggagat ggctgtcaag   1260
cgggagcagc tcaagaatgg cggcctgggc gtagtctccg acgccatctt tgaactgggc   1320
aagtcactct ctgcctttaa cctggatgac acggaagtgg ctctgctgca ggctgtgctg   1380
ctaatgtcaa cagaccgctc gggcctgctg tgtgtggaca agatcgagaa gagtcaggag   1440
gcgtacctgc tggcgttcga gcactacgtc aaccaccgca aacacaacat tccgcacttc   1500
tggcccaagc tgctgatgaa ggagagagaa gtgcagagtt cgattctgta caaggggca    1560
gcggcagaag gccggccggg cgggtcactg ggcgtccacc cggaaggaca gcagcttctc   1620
ggaatgcatg ttgttcaggg tccgcaggtc cggcagcttg agcagcagct tggtgaagcg   1680
ggaagtctcc aagggccggt tcttcagcac cagagcccga agagcccgca gcagcgtctc   1740
ctggagctgc tccaccgaag cggaattctc catgcccgag cggtctgtgg ggaagacgac   1800
agcagtgagg cggactcccc gagctcctct gaggaggaac cggaggtctg cgaggacctg   1860
gcaggcaatg cagcctctcc ctgaagcccc ccagaaggcc gatggggaag gagaaggagt   1920
gccatacctt ctcccaggcc tctgccccaa gagcaggagg tgcctgaaag ctgggagcgt   1980
gggctcagca gggctggtca cctcccatcc cgtaagacca ccttcccttc ctcagcaggc   2040
caaacatggc cagactccct tgctttttgc tgtgtagttc cctctgcctg ggatgccctt   2100
ccccctttct ctgcctggca acatcttact tgtcctttga ggccccaact caagtgtcac   2160
ctccttcccc agctccccca ggcagaaata gttgtctgtg cttccttggt tcatgcttct   2220
actgtgacac ttatctcact gttttataat tagtcgggca tgagtctgtt tcccaagcta   2280
gactgtgtct gaatcatgtc tgtatcccg                                     2309
```

<210> 16

<211> 2355

<212> DNA

<213> Homo sapiens

<400> 16

```
ccgttgcctc aacgtccaac ccttctgcag ggctgcagtc cggccacccc aagaccttgc     60
tgcagggtgc ttcggatcct gatcgtgagt cgcggggtcc actccccgcc cttagccagt    120
gcccaggggg caacagcggc gatcgcaacc tctagtttga gtcaaggtcc agtttgaatg    180
accgctctca gctggtgaag acatgaccac cctggactcc aacaacaaca caggtggcgt    240
catcacctac attggctcca gtggctcctc cccaagccgc accagccctg aatccctcta    300
tagtgacaac tccaatggca gcttccagtc cctgacccaa ggctgtccca cctacttccc    360
accatccccc actggctccc tcacccaaga cccggctcgc tcctttggga gcattccacc    420
cagcctgagt gatgacggct ccccttcttc ctcatcttcc tcgtcgtcat cctcctcctc    480
cttctataat gggagccccc ctgggagtct acaagtggcc atggaggaca gcagccgagt    540
gtcccccagc aagagcacca gcaacatcac caagctgaat ggcatggtgt tactgtgtaa    600
agtgtgtggg gacgttgcct cgggcttcca ctacggtgtg ctcgcctgcg agggctgcaa    660
gggctttttc cgtcggagca tccagcagaa catccagtac aaaaggtgtc tgaagaatga    720
gaattgctcc atcgtccgca tcaatcgcaa ccgctgccag caatgtcgct tcaagaagtg    780
tctctctgtg ggcatgtctc gagacgctgt gcgttttggg cgcatcccca aacgagagaa    840
gcagcggatg cttgctgaga tgcagagtgc catgaacctg gccaacaacc agttgagcag    900
ccagtgcccg ctggagactt cacccaccca gcaccccacc ccaggcccca tgggcccctc    960
gccacccct gctccggtcc cctcacccct ggtgggcttc tcccagtttc cacaacagct   1020
gacgcctccc agatccccaa gccctgagcc cacagtggag gatgtgatat cccaggtggc   1080
ccgggcccat cgagagatct tcacctacgc ccatgacaag ctgggcagct cacctggcaa   1140
cttcaatgcc aaccatgcat caggtagccc tccagccacc accccacatc gctgggaaaa   1200
tcagggctgc ccacctgccc ccaatgacaa caacaccttg gctgcccagc gtcataacga   1260
ggccctaaat ggtctgcgcc aggctccctc ctcctaccct cccacctggc ctcctggccc   1320
tgcacaccac agctgccacc agtccaacag caacgggcac cgtctatgcc ccacccacgt   1380
gtatgcagcc ccagaaggca aggcacctgc caacagtccc cggcagggca actcaaagaa   1440
tgttctgctg gcatgtccta tgaacatgta cccgcatgga cgcagtgggc gaacggtgca   1500
ggagatctgg gaggatttct ccatgagctt cacgcccgct gtgcgggagg tggtagagtt   1560
tgccaaacac atcccgggct tccgtgacct ttctcagcat gaccaagtca ccctgcttaa   1620
```

```
ggctggcacc tttgaggtgc tgatggtgcg ctttgcttcg ttgttcaacg tgaaggacca 1680
gacagtgatg ttcctaagcc ggaccaccta cagcctgcag gagcttggtg ccatgggcat 1740
gggagacctg ctcagtgcca tgttcgactt cagcgagaag ctcaactccc tggcgcttac 1800
cgaggaggag ctgggcctct tcaccgcggt ggtgcttgtc tctgcagacc gctcgggcat 1860
ggagaattcc gcttcggtgg agcagctcca ggagacgctg ctgcgggctc ttcgggctct 1920
ggtgctgaag aaccggccct tggagacttc ccgcttcacc aagctgctgc tcaagctgcc 1980
ggacctgcgg accctgaaca acatgcattc cgagaagctg ctgtccttcc gggtggacgc 2040
ccagtgaccc gcccggccgg ccttctgccg ctgccccctt gtacagaatc gaactctgca 2100
cttctctctc ctttacgaga cgaaaaggaa aagcaaacca gaatcttatt tatattgtta 2160
taaaatattc caagatgagc ctctggcccc ctgagccttc ttgtaaatac ctgcctccct 2220
cccccatcac cgaacttccc ctcctcccct atttaaacca ctctgtctcc cccacaaccc 2280
tcccctggcc ctctgatttg ttctgttcct gtctcaaatc caatagttca cagctaaaaa 2340
aaaaaaaaaa aaaag                                                  2355
<210> 17

<211> 4119

<212> DNA

<213> Homo sapiens


<400> 17
gaattccgtt gctgtcgcac acacacacac acacacacac acaccccaac acacacacac 60
acaccccaac acacacacac acacacacac acacacacac acacacacac acacagcggg 120
atggccgagc gccgcacgcg tagcacgccg ggactagcta tccagcctcc cagcagcctc 180
tgcgacgggc gcggtgcgta agtacctcgc cggtggtggc cgttctccgt aagatggcgg 240
accggcggcg gcagcgcgct tcgcaagaca ccgaggacga ggaatctggt gcttcgggct 300
ccgacagcgg cggctccccg ttgcggggag gcgggagctg cagcggtagc gccggaggcg 360
gcggcagcgg ctctctgcct tcacagcgcg gaggccgaac cggggccctt catctgcggc 420
gggtggagag cgggggcgcc aagagtgctg aggagtcgga gtgtgagagt gaagatggca 480
ttgaaggtga tgctgttctc tcggattatg aaagtgcaga agactcggaa ggtgaagaag 540
gtgaatacag tgaagaggaa aactccaaag tggagctgaa atcagaagct aatgatgctg 600
ttaattcttc aacaaaagaa gagaagggag aagaaaagcc tgacaccaaa agcactgtga 660
ctggagagag gcaaagtggg gacggacagg agagcacaga gcctgtggag aacaaagtgg 720
gtaaaaaggg ccctaagcat ttggatgatg atgaagatcg gaagaatcca gcatacatac 780
ctcggaaagg gctcttcttt gagcatgatc ttcgagggca aactcaggag gaggaagtca 840
gacccaaggg gcgtcagcga aagctatgga aggatgaggg tcgctgggag catgacaagt 900
tccgggaaga tgagcaggcc ccaaagtccc gacaggagct cattgctctt tatggttatg 960
acattcgctc agctcataat cctgatgaca tcaaacctcg aagaatccgg aaaccccgat 1020
atgggagtcc tccacaaaga gatccaaact ggaacggtga gcggctaaac aagtctcatc 1080
gccaccaggg tcttgggggc accctaccac caaggacatt tattaacagg aatgctgcag 1140
gtaccggccg tatgtctgca cccaggaatt attctcgatc tgggggcttc aaggaaggtc 1200
gtgctggttt taggcctgtg gaagctggtg ggcagcatgg tggccggtct ggtgagactg 1260
ttaagcatga gattagttac cggtcacggc gcctagagca gacttctgtg agggatccat 1320
ctccagaagc agatgctcca gtgcttggca gtcctgagaa ggaagaggca gcctcagagc 1380
caccagctgc tgctcctgat gctgcaccac cacccccctga taggcccatt gagaagaaat 1440
cctattcccg ggcaagaaga actcgaacca aagttggaga tgcagtcaag cttgcagagg 1500
aggtgccccc tcctcctgaa ggactgattc cagcacctcc agtcccagaa accacccccaa 1560
ctccacctac taagactggg acctgggaag ctccggtgga ttctagtaca agtggacttg 1620
agcaagatgt ggcacaacta aatatagcag aacagaattg gagtccgggg cagccttctt 1680
tcctgcaacc acgggaactt cgaggtatgc ccaaccatat acacatggga gcaggacctc 1740
cacctcagtt taaccggatg gaagaaatgg gtgtccaggg tggtcgagcc aaacgctatt 1800
catcccagcg gcaaagacct gtgccagagc cccccgcccc tccagtgcat atcagtatca 1860
tggagggaca ttactatgat ccactgcagt tccagggacc aatctatacc catggtgaca 1920
gccctgcccc gctgcctcca cagggcatgc ttgtgcagcc aggaatgaac cttccccacc 1980
caggtttaca tccccaccag acaccagctc ctctgcccaa tccaggcctc tatcccccac 2040
cagtgtccat gtctccagga cagccaccac ctcagcagtt gcttgctcct acttactttt 2100
ctgctccagg cgtcatgaac tttggtaatc ccagttaccc ttatgctcca ggggcactgc 2160
ctcccccacc accgcctcat ctgtatccta atacacaggc ccccatcacag gtatatggag 2220
gagtgaccta ctataacccc gcccagcagc aggtgcagcc aaagccctcc ccaccccgga 2280
ggactcccca gccagtcacc atcaagcccc ctccacctga ggttgtaagc aggggttcca 2340
gttaatacaa gtttctgaat attttaaatc ttaacatcat ataaaaagca gcagaggtga 2400
gaactcagaa gagaaataca gctggctatc tactaccaga agggcttcaa agatataggg 2460
tgtggctcct accagcaaac agctgaaaga ggaggacccc tgccttcctc tgaggacagg 2520
```

ctctagagag agggagaaac aagtggacct cgtcccatct tcactcttca cttgagttgg 2580
ctgtgttcgg gggagcagag agagccagac agccccaagc ttctgagtct agatacagaa 2640
gcccatgtct tctgctgttc ttcacttctg ggaaattgaa gtgtcttctg ttcccaagga 2700
agctccttcc tgtttgtttt gttttctaag atgttcattt ttaaagcctg gcttcttatc 2760
cttaatatta ttttaatttt ttctctttgt ttctgtttct tgctctctct ccctgccttt 2820
aaatgaaaca agtctagtct tctggttttc tagcccctct ggattccctt ttgactcttc 2880
cgtgcatccc agataatgga gaatgtatca gccagccttc cccaccaagt ctaaaaagac 2940
ctggcctttc acttttagtt ggcatttgtt atcctcttgt atacttgtat tcccttaact 3000
ctaaccctgt ggaagcatgg ctgtctgcac agagggtccc attgtgcaga aaagctcaga 3060
gtaggtgggt aggagccctt ctctttgact taggttttta ggagtctgag catccatcaa 3120
tacctgtact atgatgggct tctgttctct gctgagggcc aataccctac tgtggggaga 3180
gatggcacac cagatgcttt tgtgagaaag ggatggtgga gtgagagcct ttgcctttag 3240
gggtgtgtat tcacatagtc ctcagggctc agtcttttga ggtaagtgga attagagggc 3300
cttgcttctc ttctttccat tcttcttgct acacccettt tccagttgct gtggaccaat 3360
gcatctcttt aaaggcaaat attatccagc aagcagtcta ccctgtcctt tgcaattgct 3420
cttctccacg tctttcctgc tacaagtgtt ttagatgtta ctaccttatt ttccccgaat 3480
tctatttttg tccttgcaga cagaatataa aaactcctgg gcttaaggcc taaggaagcc 3540
agtcaccttc tgggcaaggg ctcctatctt tcctccctat ccatggcact aaaccacttc 3600
tctgctgcct ctgtggaaga gattcctatt actgcagtac atacgtctgc caggggtaac 3660
ctggccactg tccctgtcct tctacagaac ctgagggcaa agatggtggc tgtgtctctc 3720
cccggtaatg tcactgtttt tattccttcc atctagcagc tggcctaatc actctgagtc 3780
acaggtgtgg gatggagagt ggggagaggc acttaatctg taacccccaa ggaggaaata 3840
actaagagat tcttctaggg gtagctggtg gttgtgcctt ttgtaggctg ttccctttgc 3900
cttaaacctg aagatgtctc ctcaagcctg tgggcagcat gcccagattc ccagacctta 3960
agacactgtg agagttgtct ctgttggtcc actgtgttta gttgcaagga tttttccatg 4020
tgtggtggtg tttttgtta ctgttttaaa gggtgcccat ttgtgatcag cattgtgact 4080
tggagataat aaaatttaga ctataaactt gaaaaaaa 4119

<210> 18

<211> 2653

<212> DNA

<213> Homo sapiens

<400> 18

gagcgcggct ggagtttgct gctgccgctg tgcagtttgt tcaggggctt gtggcggtga 60
gtccgagagg ctgcgtgtga gagacgtgag aaggatcctg cactgaggag gtggaaagaa 120
gaggattgct cgaggaggcc tggggtctgt gagacagcgg agctgggtga aggctgcggg 180
ttccggcgag gcctgagctg tgctgtcgtc atgcctcaaa cccgatccca ggcacaggct 240
acaatcagtt ttccaaaaag gaagctgtct cgggcattga acaaagctaa aaactccagt 300
gatgccaaac tagaaccaac aaatgtccaa accgtaacct gttctcctcg tgtaaaagcc 360
ctgcctctca gccccaggaa acgtctgggc gatgacaacc tatgcaacac tccccattta 420
cctccttgtt ctccaccaaa gcaaggcaag aaagagaatg gtccccctca ctcacataca 480
cttaagggac gaagattggt atttgacaat cagctgacaa ttaagtctcc tagcaaaaga 540
gaactagcca aagttcacca aaacaaaata ctttcttcag ttagaaaaag tcaagagatc 600
acaacaaatt ctgagcagag atgtccactg aagaaagaat ctgcatgtgt gagactattc 660
aagcaagaag gcacttgcta ccagcaagca aagctggtcc tgaacacagc tgtcccagat 720
cggctgcctg ccagggaaag ggagatggat gtcatcagga atttcttgag ggaacacatc 780
tgtgggaaaa aagctggaag cctttacctt tctggtgctc ctggaactgg aaaaactgcc 840
tgcttaagcc ggattctgca agacctcaag aaggaactga aaggctttaa aactatcatg 900
ctgaattgca tgtccttgag gactgcccag gctgtattcc cagctattgc tcaggagatt 960
tgtcaggaag aggtatccag gccagctggg aaggacatga tgaggaaatt ggaaaaacat 1020
atgactgcag agaagggccc catgattgtg ttggtattgg acgagatgga tcaactggac 1080
agcaaaggcc aggatgtatt gtacacgcta tttgaatggc catggctaag caattctcac 1140
ttggtgctga ttggtattgc taataccctg gatctcacag atagaattct acctaggctt 1200
caagctagag aaaaatgtaa gccacagctg ttgaacttcc caccttatac cagaaatcag 1260
atagtcacta ttttgcaaga tcgacttaat caggtatcta gagatcaggt tctggacaat 1320
gctgcagttc aattctgtgc ccgcaaagtc tctgctgttt caggagatgt tcgcaaagca 1380
ctggatgttt gcaggagagc tattgaaatt gtagagtcag atgtcaaaag ccagactatt 1440
ctcaaaccac tgtctgaatg taaatcacct tctgagcctc tgattcccaa gagggttggt 1500
cttattcaca tatcccaagt catctcagaa gttgatggta acaggatgac cttgagccaa 1560
gagggagcac aagattcctt ccctcttcag cagaagatct tggtttgctc tttgatgctc 1620
ttgatcaggc agttgaaaat caaagaggtc actctgggga agttatatga agcctacagt 1680

aaagtctgtc gcaaacagca ggtggcggct gtggaccagt cagagtgttt gtcactttca 1740
gggctcttgg aagccagggg cattttagga ttaaagagaa acaaggaaac ccgtttgaca 1800
aaggtgtttt tcaagattga agagaaagaa atagaacatg ctctgaaaga taaagcttta 1860
attggaaata tcttagctac tggattgcct taaattcttc tcttacaccc cacccgaaag 1920
tattcagctg gcatttagag agctacagtc ttcattttag tgctttacac attcgggcct 1980
gaaaacaaat atgacctttt ttacttgaag ccaatgaatt ttaatctata gattctttaa 2040
tattagcaca gaataatatc tttgggtctt actattttta cccataaaag tgaccaggta 2100
gacccttttt aattacattc actacttcta ccacttgtgt atctctagcc aatgtgcttg 2160
caagtgtaca gatctgtgta gaggaatgtg tgtatattta cctcttcgtt tgctcaaaca 2220
tgagtgggta tttttttgtt tgtttttttt gttgttgttg tttttgaggc gcgtctcacc 2280
ctgttgccca ggctggagtg caatggcgcg ttctctgctc actacagcac ccgcttccca 2340
ggttgaagtg attctcttgc ctcagcctcc cgagtagctg ggattacagg tgcccaccac 2400
cgcgcccagc taattttttta atttttagta gagacagggt tttaccatgt tggccaggct 2460
ggtcttgaac tcctgaccct caagtgatct gcccaccttg gcctccctaa gtgctgggat 2520
tataggcgtg agccaccatg ctcagccatt aaggtatttt gttaagaact ttaagtttag 2580
ggtaagaaga atgaaaatga tccagaaaaa tgcaagcaag tccacatgga gatttggagg 2640
acactggtta aag 2653

<210> 19

<211> 2907

<212> DNA

<213> Homo sapiens

<400> 19
gccatctggg cccaggcccc atgccccgag gaggggtggt ctgaagccca ccagagcccc 60
ctgccagact gtctgcctcc cttctgactg tggccgcttg gcatggccag caacagcagc 120
tcctgcccga cacctggggg cgggcacctc aatgggtacc cggtgcctcc ctacgccttc 180
ttcttccccc ctatgctggg tggactctcc ccgccaggcg ctctgaccac tctccagcac 240
cagcttccag ttagtggata tagcacacca tccccagcca ccattgagac ccagagcagc 300
agttctgaag agatagtgcc cagccctccc tcgccacccc ctctaccccg catctacaag 360
ccttgctttg tctgtcagga caagtcctca ggctaccact atggggtcag cgcctgtgag 420
ggctgcaagg gcttcttccg ccgcagcatc cagaagaaca tggtgtacac gtgtcaccgg 480
gacaagaact gcatcatcaa caaggtgacc cggaaccgct gccagtactg ccgactgcag 540
aagtgctttg aagtgggcat gtccaaggag tctgtgagaa acgaccgaaa caagaagaag 600
aaggaggtgc ccaagcccga gtgctctgag agctacacgc tgacgccgga ggtgggggag 660
ctcattgaga aggtgcgcaa agcgcaccag gaaaccttcc ctgccctctg ccagctgggc 720
aaatacacta cgaacaacag ctcagaacaa cgtgtctctc tggacattga cctctgggac 780
aagttcagtg aactctccac caagtgcatc attaagactg tggagttcgc caagcagctg 840
cccggcttca ccaccctcac catcgccgac cagatcaccc tcctcaaggc tgcctgcctg 900
gacatcctga tcctgcggat ctgcacgcgg tacacgcccg agcaggacac catgaccttc 960
tcggacgggc tgaccctgaa ccggacccag atgcacaacg ctggcttcgg cccccctcacc 1020
gacctggtct ttgccttcgc caaccagctg ctgcccctgg agatggatga tgcggagacg 1080
gggctgctca gcgccatctg cctcatctgc ggagaccgcc aggacctgga gcagccggac 1140
cgggtggaca tgctgcagga gccgctgctg gaggcgctaa aggtctacgt gcggaagcgg 1200
aggcccagcc gcccccacat gttccccaag atgctaatga agattactga cctgcgaagc 1260
atcagcgcca aggggggctga gcgggtgatc acgctgaaga tggagatccc gggctccatg 1320
ccgcctctca tccaggaaat gttggagaac tcagagggcc tggacactct gagcggacag 1380
ccgggggggtg gggggcggga cgggggtggc ctggcccccc cgccaggcag ctgtagcccc 1440
agcctcagcc ccagctccaa cagaagcagc ccggccaccc actccccgtg accgcccacg 1500
ccacatggac acagccctcg ccctccgccc cggcttttct ctgcctttct accgaccatg 1560
tgaccccgca ccagccctgc ccccacctgc cctcccgggc agtactgggg accttccctg 1620
ggggacgggg agggaggagg cagcgactcc ttggacagag gcctgggccc tcagtggact 1680
gcctgctccc acagcctggg ctgacgtcag aggccgaggc caggaactga gtgaggcccc 1740
tggtcctggg tctcaggatg ggtcctgggg gcctcgtgtt catcaagaca cccctctgcc 1800
cagctcacca catcttcatc accagcaaac gccaggactt ggctccccca tcctcagaac 1860
tcacaagcca ttgctcccca gctggggaac ctcaacctcc cccctgcctc ggttggtgac 1920
agaggggggtg ggacaggggc gggggggttcc ccctgtacat accctgccat accaacccca 1980
ggtattaatt ctcgctggtt ttgtttttat tttaattttt ttgttttgat tttttttaata 2040
agaattttca ttttaagcac atttatactg aaggaatttg tgctgtgtat tggggggagc 2100
tggatccaga gctggagggg gtgggtccgg gggagggagt ggctcggaag gggcccccac 2160
tctcctttca tgtccctgtg ccccccagtt ctcctcctca gccttttcct cctcagtttt 2220
ctctttaaaa ctgtgaagta ctaactttcc aaggcctgcc ttcccctccc tcccactgga 2280

gaagccgcca gcccctttct ccctctgcct gaccactggg tgtggacggt gtggggcagc 2340
cctgaaagga caggctcctg gccttggcac ttgcctgcac ccaccatgag gcatggagca 2400
gggcagagca agggccccgg gacagagttt tcccagacct ggctcctcgg cagagctgcc 2460
tcccgtcagg gcccacatca tctaggctcc ccagccccca ctgtgaaggg gctggccagg 2520
ggcccgagct gcccccaccc ccggcctcag ccaccagcac ccccataggg cccccagaca 2580
ccacacacat gcgcgtgcgc acacacacaa acacacacac actggacagt agatgggccg 2640
acacacactt ggcccgagtt cctccatttc cctggcctgc cccccacccc caacctgtcc 2700
cacccccgtg ccccctcctt accccgcagg acgggcctac agggggtct cccctcaccc 2760
ctgcaccccc agctggggga gctggctctg ccccgacctc cttcaccagg ggttggggcc 2820
ccttcccctg gagcccgtgg gtgcacctgt tactgttggg ctttccactg agatctactg 2880
gataaagaat aaagttctat ttattct 2907

<210> 20

<211> 2096

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (23)..(23)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (27)..(27)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (80)..(80)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (120)..(120)

<223> n=a, c, g or t

<400> 20

agatgtttaa aaatactttg atnctcngtt tccacctctc ttaaattgtc tttccctatg 60
ttaaatatac agtcatcacn ttgctgaaaa aagttcgcaa tgagaacaat catctaaaan 120
tggctgtaac taggtcaggc gcggttgctc atgcctgtaa tcccaccact ttgggaggcc 180
gaggcaattg gatcacctga ggtcaggatt ttgagaccag cttgaccaac atggtggaat 240
cccatctcta ctaaaaatac aaaaaattag ccgggtgtgg tggcacaccc ctgtaatccc 300
acctactcag gaggctgagg caggaaaatc ccttgaaccc aggaggcaaa ggttgcattg 360
agccgaaata acaccactgc actccagcct ggacgataga gtgagacccc atctcaaaaa 420
aagagcagct gtgacaaatg cctgtattga attgcaggtc agtcttccac ctccactacc 480

ggtgccaaaa aaagggctgc cccaaaagga actaaaaggg atccagcttt gaattctggt 540
gtctctcaaa agcctgatcc tgccaaaacc aagaatcgcc gcaaaaggaa gccatccact 600
tctgatgatt ctgactctaa ttttgagaaa attgtttcga aagcagtcac aagcaaggtg 660
agtgttgatc ctagtcagtc cttttgctgt agatgttctg aaacacgtaa ctaagccatt 720
gttcttaaaa atttggcata tctttaagaa aattaactct catattctgt tagcttttac 780
tgtacatatt tagttttaac aaagttaaat atgccactta tttggccaat ggaagagttg 840
gccttagatc tgcttcttat tacttggtag aaaatagaaa actccttgaa tatagtgtct 900
tgatacattt ttttacatta caattatgtt gtcagattta caatgtgcaa gttacctggg 960
cttttctctt ttagaaatcc aagggggaga gtgatgactt ccatatggac tttgactcag 1020
ctgtggctcc tcgggcaaaa tctgtacggg caaagaaacc tataaagtac ctggaagagt 1080
cagatgaaga tgatctgttt taaaatgtga ggcgattatt ttaagtaatt atcttaccaa 1140
gcccaagact ggttttaaag ttacctgaag ctcttaactt cctcccctct gaatttagtt 1200
tggggaaggt gtttttagta caagacatca aagtgaagta aagcccaagt gttctttagc 1260
tttttataat actgtataaa tagtgaccat ctcatgggca ttgttttctt ctctgctttg 1320
tctgtgtttt gagtctgctt cttttgtctt taaaacctga tttttaagtt cttctgaact 1380
gtagaaatag ctatctgatc acttcagcgt aaagcagtgt gtttattaac catccactaa 1440
gctaaaacta gagcagtttg atttaaaagt gtcactcttc ctccttttct actttcagta 1500
gatatgagat agagcataat tatctgtttt atcttagttt tatacataat ttaccatcag 1560
atagaacttt atggttctag tacagatact ctactacact cagcctctta tgtgccaagt 1620
ttttctttaa gcaatgagaa attgctcatg ttcttcatct tctcaaatca tcagaggccg 1680
aagaaaaaca ctttggctgt gtctataact tgacacagtc aatagaatga agaaaattag 1740
agtagttatg tgattatttc agctcttgac ctgtcccctc tggctgcctc tgagtctgaa 1800
tctcccaaag agagaaacca atttctaaga ggactggatt gcagaagact cggggacaac 1860
atttgatcca agatcttaaa tgttatattg ataaccatgc tcagcaatga gctattagat 1920
tcattttggg aaatctccat aatttcaatt tgtaaacttt gttaagacct gtctacattg 1980
ttatatgtgt gtgacttgag taatgttatc aacgtttttg taaatattta ctatgttttt 2040
ctattagcta aattccaaca attttgtact ttaataaaat gttctaaaca ttgaaa 2096

<210> 21

<211> 2160

<212> DNA

<213> Homo sapiens

<400> 21
agccccctgc ccctcgccgc cccccgccgc ctgcctgggc cgggccgagg atgcggcgca 60
gcgcctcggc ggccaggctt gctcccctcc ggcacgcctg ctaacttccc ccgctacgtc 120
cccgttcgcc cgccgggccg ccccgtctcc ccgcggcctc cgggtccggg tcctccagga 180
cggccaggcc gtgccgccgt gtgccctccg ccgctcgccc gcgcgccgcg cgctccccgc 240
ctgcgcccag cgccccgcgc ccgcgcccca gtcctcgggc ggtccatgct gcccctctgc 300
ctcgtggccg ccctgctgct ggccgccggg cccgggccga gcctgggcga cgaagccatc 360
cactgcccgc cctgctccga ggagaagctg gcgcgctgcc gcccccccgt gggctgcgag 420
gagctggtgc gagaggcggg ctgcggctgt tgcgccactt gcgccctggg cttggggatg 480
ccctgcgggg tgtacacccc ccgttgcggc tcgggcctgc gctgctaccc gccccgaggg 540
gtggagaagc ccctgcacac actgatgcac gggcaaggcg tgtgcatgga gctggcggag 600
atcgaggcca tccaggaaag cctgcagccc tctgacaagg acgagggtga ccaccccaac 660
aacagcttca gcccctgtag cgcccatgac cgcaggtgcc tgcagaagca cttcgccaaa 720
attcgagacc ggagcaccag tgggggcaag atgaaggtca atggggcgcc ccgggaggat 780
gcccggcctg tgccccaggg ctcctgccag agcgagctgc accgggcgct ggagcggctg 840
gccgcttcac agagccgcac ccacgaggac ctctacttca tccccatccc caactgcgac 900
cgcaacggca acttccaccc caagcagtgt cacccagctc tggatgggca gcgtggcaag 960
tgctggtgtg tggaccggaa gacgggggtg aagcttccgg ggggcctgga gccaaagggg 1020
gagctggact gccaccagct ggctgacagc tttcgagagt gaggcctgcc agcaggccag 1080
ggactcagcg tcccctgcta ctcctgtgct ctggaggctg cagagctgac ccagagtgga 1140
gtctgagtct gagtcctgtc tctgcctgcg gcccagaagt ttccctcaaa tgcgcgtgtg 1200
cacgtgtgcg tgtgcgtgcg tgtgtgtgtg tttgtgagca tgggtgtgcc cttggggtaa 1260
gccagagcct ggggtgttct ctttggtgtt acacagccca agaggactga gactggcact 1320
tagcccaaga ggtctgagcc ctggtgtgtt tccagatcga tcctggattc actcactcac 1380
tcattccttc actcatccag ccacctaaaa acatttactg accatgtact acgtgccagc 1440
tctagttttc agccttggga ggttttattc tgacttcctc tgattttggc atgtggagac 1500
actcctataa ggagagttca agcctgtggg agtagaaaaa tctcattccc agagtcagag 1560
gagaagagac atgtaccttg accatcgtcc ttcctctcaa gctagcccag agggtgggag 1620
cctaaggaag cgtggggtag cagatggagt aatggtcacg aggtccagac ccactcccaa 1680

agctcagact tgccaggctc cctttctctt cttccccagg tccttccttt aggtctggtt 1740
gttgcaccat ctgcttggtt ggctggcagc tgagagccct gctgtgggag agcgaagggg 1800
gtcaaaggaa gacttgaagc acagagggct agggaggtgg ggtacatttc tctgagcagt 1860
cagggtggga agaaagaatg caagagtgga ctgaatgtgc ctaatggaga agacccacgt 1920
gctaggggat gaggggcttc ctgggtcctg ttcccctacc ccatttgtgg tcacagccat 1980
gaagtcaccg ggatgaacct atccttccag tggctcgctc cctgtagctc tgcctccctc 2040
tccatatctc cttcccctac acctccctcc ccacacctcc ctactcccct gggcatcttc 2100
tggcttgact ggatggaagg agacttagga acctaccagt tggccatgat gtcttttctt 2160

<210> 22

<211> 2215

<212> DNA

<213> Homo sapiens

<400> 22
ctgcagggag ccatgattgc accactgcac tccagcctgg gcaacagagt gagaccatgt 60
ctcaagaaaa aaaaaaaaga aagaaaccac tgctctaggc taaatcccag ccagagttgg 120
agccacccag ctaaactggc ctgttttccc tcatttcctt ccccgaaggt atgcctgtgt 180
caagatgagg tcacggacga ttacatcgga gacaacacca cagtggacta cactttgttc 240
gagtctttgt gctccaagaa ggacgtgcgg aactttaaag cctggttcct ccctatcatg 300
tactccatca tttgtttcgt gggcctactg ggcaatgggc tggtcgtgtt gacctatatc 360
tatttcaaga ggctcaagac catgaccgat acctacctgc tcaacctggc ggtggcagac 420
atcctcttcc tcctgaccct tcccttctgg gcctacagcg cggccaagtc ctgggtcttc 480
ggtgtccact tttgcaagct catctttgcc atctacaaga tgagcttctt cagtggcatg 540
ctcctacttc tttgcatcag cattgaccgc tacgtggcca tcgtccaggc tgtctcagct 600
caccgccacc gtgcccgcgt ccttctcatc agcaagctgt cctgtgtggg catctggata 660
ctagccacag tgctctccat cccagagctc ctgtacagtg acctccagag gagcagcagt 720
gagcaagcga tgcgatgctc tctcatcaca gagcatgtgg aggcctttat caccatccag 780
gtggcccaga tggtgatcgg ctttctggtc cccctgctgg ccatgagctt ctgttacctt 840
gtcatcatcc gcaccctgct ccaggcacgc aactttgagc gcaacaaggc catcaaggtg 900
atcatcgctg tggtcgtggt cttcatagtc ttccagctgc cctacaatgg ggtggtcctg 960
gcccagacgg tggccaactt caacatcacc agtagcacct gtgagctcag taagcaactc 1020
aacatcgcct acgacgtcac ctacagcctg gcctgcgtcc gctgctgcgt caacccttc 1080
ttgtacgcct tcatcggcgt caagttccgc aacgatctct tcaagctctt caaggacctg 1140
ggctgcctca gccaggagca gctccggcag tggtcttcct gtcggcacat ccggcgctcc 1200
tccatgagtg tggaggccga gaccaccacc accttctccc cataggcgac tcttctgcct 1260
ggactagagg gacctctccc agggtccctg ggtggggat agggagcaga tgcaatgact 1320
caggacatcc ccccgccaaa agctgctcag ggaaaagcag ctctcccctc agagtgcaag 1380
ccctgctcca gaagttagct tcaccccaat cccagctacc tcaaccaatg ccgaaaaaga 1440
cagggctgat aagctaacac cagacagaca acactgggaa acagaggcta ttgtccccta 1500
aaccaaaaac tgaaagtgaa agtccagaaa ctgttcccac ctgctggagt gaaggggcca 1560
aggagggtga gtgcaagggg cgtgggagtg gcctgaagag tcctctgaat gaaccttctg 1620
gcctcccaca gactcaaatg ctcagaccag ctcttccgaa aaccaggcct tatctccaag 1680
accagagata gtggggagac ttcttggctt ggtgaggaaa agcggacatc agctggtcaa 1740
acaaactctc tgaacccctc cctccatcgt ttcttcact gtcctccaag ccagcgggaa 1800
tggcagctgc cacgccgccc taaaagcaca ctcatcccct cacttgccgc gtcgccctcc 1860
caggctctca acaggggaga gtgtggtgtt tcctgcaggc caggccagct gcctccgcgt 1920
gatcaaagcc acactctggg ctccagagtg gggatgacat gcactcagct cttggctcca 1980
ctgggatggg aggagaggac aagggaaatg tcaggggcgg ggagggtgac agtggccgcc 2040
caaggccacg agcttgttct ttgttctttg tcacagggac tgaaacctc tcctcatgtt 2100
ctgctttcga ttcgttaaga gagcaacatt ttacccacac acagataaag ttttcccttg 2160
aggaaacaac agctttaaaa gaaaaagaa aaaaaagct tggtaagtca agtag 2215

<210> 23

<211> 958

<212> DNA

<213> Homo sapiens

<400> 23
ggggccggac gcgaggggcg gggcgagcgc gggacaaagg gaagcgaagc cggagctgcg 60
ggcgcttttt ctgcccgcgg tgtctcagat tcattcttaa ggaactgaga acttaatctt 120
ccaaaatgtc aaaaagacca tcttatgccc cacctcccac cccagctcct gcaacacaaa 180
tgcccagcac accagggttt gtgggataca atccatacag tcatctcgcc tacaacaact 240
acaggctggg agggaacccg agcaccaaca gccgggtcac ggcatcctct ggtatcacga 300
ttccaaaacc cccaaagcca ccagataagc cgctgatgcc ctacatgagg tacagcagaa 360
aggtctggga ccaagtaaag gcttccaacc ctgacctaaa gttgtgggag attggcaaga 420
ttattggtgg catgtggcga gatctcactg atgaagaaaa acaagaatat ttaaacgaat 480
acgaagcaga aaagatagag tacaatgaat ctatgaaggc ctatcataat tccccccgcgt 540
accttgctta cataaatgca aaaagtcgtg cagaagctgc tttagaggaa gaaagtcgac 600
agagacaatc tcgcatggag aaaggagaac cgtacatgag cattcagcct gctgaagatc 660
cagatgatta tgatgatggc ttttcaatga agcatacagc caccgcccgt ttccagagaa 720
accaccgcct catcagtgaa attcttagtg agagtgtggt gccagacgtt cggtcagttg 780
tcacaacagc tagaatgcag gtcctcaaac ggcaggtcca gtccttaatg gttcatcagc 840
gaaaactaga agctgaactt cttcaaatag aggaacgaca ccaggagaag aagaggaaat 900
tcctggaaag cacagattca tttaacaatg aacttaaaag gttgtgcggt ctgaaagt 958
<210> 24

<211> 6483

<212> DNA

<213> Homo sapiens

<400> 24
aagcttctaa ttgcagttca accacctgtt acatatcttc aggaaaaaat cacaacctct 60
caacttcaac ttcctcttct ataaattaga aataacaata accacacctg taaccccagc 120
actttgggag gccaaggcag gcagatcaag aggtgaggag attgagacca tcctggctaa 180
catgatgaaa ccctgtctct accaaaaaga caaaaaatta gccaggtatg gtggcacaca 240
cctgtagtcc cagctactcg ggaggctgag gcaggagaat ggcgtgaacc cgggaggtgg 300
agcttgcagt gagccgagat ggcgccactg cactccagcc tgggcgacag agcaagcctc 360
cgtctaaaaa aaaaaaaaga aagaaagaaa gaaagaaaga aaagaaataa taataaccac 420
cattcctatc tcaacagctt gttctagaaa tttttaaagc acagtatcac aaacagcact 480
acataattgt aaaacatgta tgaatatata catccaaaca acagcaatgt catagcctat 540
gggtagatat aatcttatac aatgtaccaa aatcccaatt tacttcacta gacaaactgt 600
tataccaaat tctgtacaca gtatatccaa gaaaatgtgt tgtttttatt gagaaactga 660
acctagcttg ggaacacatg tgcacagtct agttcataat atttggtgca agtatcattc 720
tctaatatag atttacattt ttgcaagcaa atttttactt gcaatcgtaa catatccaaa 780
ttttcccttt ttactcaatc agaacttagt gtaaagtact acaagttagt tcttcggatt 840
tcatgctaag aaaataatgc agattttctg cattattatg gtcttcacag aaaccttaac 900
tatgatgaat ttaaaagtgc aaaataatcc aggataactt tatgatttca catttttttaa 960
tgttaaaaat aatgccatca ttaattagaa aattctaaaa tcattacttc cactttctta 102
ggcaaaatat caatatactc tcatttgcca aataaattaa aagatctcct acaaacacaa 1080
tctcctaaat tgtggtttta tggctttaat gttttatgtg tggcaactat tgatgctagt 1140
taaaatttta gaaactcttt ctttttgatt ccctacagtt gtctacaaga accttattgt 1200
agcatgatcc tgccagactt tatactattt gttgctccaa ttaaaactgt ttaaaacatg 1260
aatttgaaaa atcttatttt aactataatt ttgtagctga aacttttttt tctaaacttt 1320
gcaaacattc tatgcaacct gaattagtgc tgagaaaatt ggatcttaat ggttgctcaa 1380
tgttcttcaa caggtgaaaa gcataataaa acatgctcat ctgaactcca cccattttca 1440
atttcaacat agcatacctc gtgtttattc ttagggcaaa ttcaaaattg tacatattag 1500
gattggttat tactgaagat aatttatgca atcataagcc aaagatgcta agttggcaaa 1560
aagaaaacaa tgtaagtaag caaactctaa cacatgtgga cacaccctct cagtatataa 1620
aggcttgtca ctgtccttgg tagcaggcac tccctgggct aaacagcatc accatgtctg 1680
ttcgatacag ctcaagcaag cactactctt cctcccgcag tggaggagga ggaggaggag 1740
gaggatgtgg aggaggagga ggagtgtcat ccctaagaat ttctagcagc aaaggctccc 1800
ttggtggagg atttagctca ggggggttca gtggtggctc ttttagccgt gggagctctg 1860
gtgggggatg ctttgggggc tcatcaggtg gctatggagg attaggaggt tttggtggag 1920
gtagctttca tggaagctat ggaagtagca gctttggtgg gagttatgga ggcagctttg 1980
gaggggggcaa tttcggaggt ggcagctttg gtgggggcag ctttggtgga ggcggctttg 2040
gtggaggcgg ctttggagga ggctttggtg gtggatttgg aggagatggt ggccttctct 2100
ctggaaatga aaaagtaacc atgcagaatc tgaatgaccg cctggcttcc tacttggaca 2160
aagttcgggc tctggaagaa tcaaactatg agctggaagg caaaatcaag gagtggtatg 2220
aaaagcatgg caactcacat caggggggagc ctcgtgacta cagcaaatac tacaaaacca 2280

tcgatgacct taaaaatcag gtaagaggta tttttaaatc cagctttaag tatcttgtcc 2340
atgtaatcca gacagatgaa tcttaaatta agcacaatgt ggctgttcac tatgcttacc 2400
catgttactt tcttccttca aaaataaccc agtctcatca aagataaaca tctgtgaaac 2460
tatggtcatg gcaatcttca tccagcaagt gtgctacttg tcttaagagg atgggagatt 2520
tactaagcac ttttgaggtt ttaatgagca tacaatgagt ccacagttaa aatatgctag 2580
gctatttaca aatgtagaaa ctgaaaaaaa aaatcatgat atgaatcaga acaaaatgtt 2640
attcagactg ataacaagcc atattcagta ccaacatggc aagaaaaata aattttccag 2700
tatgaaaatg ggacactgct tgcttctaag gaatttctga attgtaccta ttgtgtacca 2760
gttcagagtg tatttattta ttagtattta tcatgagtta aacaaatgca ggtgtgagtc 2820
agccaaagca tggctgaaat acatggaaat cacatagtct aaaagaggag ggcacactta 2880
caggaataca tctatataat tccagttagt tttcagaaag gaataattcg tgtacagaaa 2940
tacaagactg gagaaattcc aagagaacaa ataattcaaa gttaagtata tgggtaagcc 3000
tgcaatattt catatttaaa ataaaaaatt ttcccaagat tttgtaagag aacaacataa 3060
aagtgcagag tgcatctatg tcactacaaa agccatatct gcatctgacc tcttctcaaa 3120
taactgtgcc tctccctcca gattctcaac ctaacaactg ataatgccaa catcctgctt 3180
cagatcgaca atgccaggct ggcagctgat gacttcaggc tgaagtaagt taagtgatcg 3240
ttgtataata ctatcacaac gaatacatca gtggttttta acaatgactt gggatgccct 3300
caataacatt tacatttttc tgaattcacc caaagttaaa tagtattgga gttatctgag 3360
aaattttcca tgtcagtgtt acctttttgg caatattaaa ggaagaaaat gcatattaaa 3420
gtaactgcta aggtttttc cattaaacca ctattacttc taagagaact gtacatgaca 3480
aatattgcca ttacatgaga tcaactatgt agttgctttt taaatagtct ctgcccagat 3540
acatctcccc tatataagtt ataaccagta ttgatatcat gcttgtttca ggtatgagaa 3600
tgaggtagct ctgcgccaga gcgtggaggc tgacatcaac ggcctgcgta gggtgctgga 3660
tgagctgacc ctgaccaagg ctgacctgga gatgcaaatt gagagcctga ctgaagagct 3720
ggcctatctg aagaagaacc acgaggaggt gacacaaaag ttatactttt cccagccaaa 3780
agagagttca ttatggtcct cgtgtagcca ataaatcttt ctgttcctca aacaggaaat 3840
gaaagacctt cgaaatgtgt ccactggtga tgtgaatgtg gaaatgaatg ctgccccggg 3900
tgttgatctg actcaacttc tgaataacat gagaagccaa tatgaacaac ttgctgaaca 3960
aaaccgcaaa gatgctgaag cctggttcaa tgaaaaggta aagtaatctt ccttatagtg 4020
aaactcatgg aggttttatc atttcagaat ttcctcaccc ttttccttgt ttttaatact 4080
ctagagcaag gaactgacta cagaaattga taataacatt gaacagatat ccagctataa 4140
atctgagatt actgaattga gacgtaatgt acaagctctg gagatagaac tacagtccca 4200
actggccttg gtatgttaac tctcatgaaa tgacttcaac tttatcatac aaagtttcat 4260
gctcacctaa gaatatgcaa tgcaacaaaa aaatgcagag ttggaggtaa gaaagagaaa 4320
acaaagtgaa gctcatgtta atggaggaaa agtactacta gtgttgatct aaaagtgctg 4380
aaactgaaat ggtgccatta aacatacaac aaattctgtt catttcttta ttcttctata 4440
taatgcctta ctaaataatc aaataagcgt caccatactc aactgaacaa ggaagtcact 4500
aagccacaaa aaaatccgtt tcagaaacaa tccctggaag cctccttggc agaaacagaa 4560
ggtcgctact gtgtgcagct ctcacagatt cacgcccaga tatccgctct ggaagaacag 4620
ttgcaacaga ttcgagctga aaccgagtgc cagaatactg aataccaaca actcctggat 4680
attaagatcc gactggagaa tgaaattcaa acctaccgca gcctgctaga aggagaggga 4740
aggtaaatta taacatgaaa agttatccca gtttctttta ttcaatattc cagatagcaa 4800
ggcttatcta aaccccaaga agatgccaga gaatgagagg aagggaggag agagggtaga 4860
gtacagaaaa aggagtacgc aaccgcaatc tcactttctc atgaatttgg cccaaaatga 4920
ttcttaagag ttctgtgaac ttaacattgt tttcaaagga tgggttttaa aatatatacc 4980
tggcagggtt ttattttttc aacacgtttt gcttattttc taaattaacg gcaactggaa 5040
agctacccac cgttttccaa cgttagagat aaccgaatgt gacctcaccc cgtttagttc 5100
cggaggcggc ggacgcggcg gcggaagttt cggcggcggc tacggcggcg gaagctccgg 5160
cggcggaagc tccggcggcg gctacggcgg cggccacggc ggcagttccg gcggcggcta 5220
cggaggcgga agctccggcg gcggaagctc cggcggcggc tacgggggcg gaagctccag 5280
cggcggccac ggcggcggaa gctccagcgg cggccacggc ggcagttcca gcggcggcta 5340
cggtggtggc agttccggcg gcggcggcgg cggctacggg ggcggcagct ccggcggcgg 5400
cagcagctcc ggcggcggat acggcggcgg cagctccagc ggaggccaca agtcctcctc 5460
ttccgggtcc gtgggcgagt cttcatctaa gggaccaagg tcagcagaaa ctagctgggg 5520
taatctagaa ttagttttaa cttcctgtga tggttttttt gcgctttaag ctctagagtt 5580
gttttaaaaa attaaaaatc ttagagacgg ttccgtttgc atttgttcac aaactactct 5640
taacaccagc cgtgaaaaat ggcatgatca aaatgtcata ccttaagcat tttttgggc 5700
ttaacaatgt aaagttgaaa tttccttctt tttacaatat ttgcttgtta attactaagg 5760
atccctacag actgtttaaa atttttttc catcattcac acagatacta acaaaaccag 5820
agtaatcaag acaattattg aagaggtggc gcccgacggt agagttcttt catctatggt 5880
tgaatcagaa accaagaaac actactatta aactgcatca agaggaaaga gtctcccttc 5940
acacagacca ttatttacag atgcatggaa aacaaagtct ccaagaaaac acttctgtct 6000
tgatggtcta tggaaataga ccttgaaaat aaggtgtcta caaggtgttt tgtggtttct 6060
gtatttcttc ttttcacttt accacaaagt gttctttaat ggaaagaaaa acaactttgt 6120
gttctcattt actaatgaat ttcaataaac tttcttactg atgcaaacta tcccaatttg 6180
tcagaattta tctttactta agtacataat actctttaaa attaaagatt agtaacccat 6240

agcagttgaa ggttgatgta tccagaaatt cggaagacag aactattgtc atgccttttc 6300
taagtttttt aatcatgtat gttcagacca ccgtcagtaa attcactgag taaagtctgt 6360
aaatccccaa tattactctt taagatacac aatatgtgga aggctcccag ctctctggct 6420
ttaaattatt tcaatcctgg aaattctgga atatctcaaa tataaccccc aaaataataa 6480
taa 6483

<210> 25

<211> 1871

<212> DNA

<213> Homo sapiens

<400> 25
agttgtggcc accttcccca ggccatggat ctctccaaca acaccatgtc actctcagtg 60
cgcacccccg gactgtcccg gcggctctcc tcgcagagtg tgataggcag acccaggggc 120
atgtctgctt ccagtgttgg aagtggttat gggggaagtg cctttggctt tggagccagc 180
tgtgggggag gcttttctgc tgcttccatg tttggttcta gttccggctt tggggggtggc 240
tccggaagtt ccatggcagg aggactgggt gctggttatg ggagagccct gggtggaggt 300
agctttggag ggctggggat gggatttggg ggcagcccag gaggtggctc tctaggtatt 360
ctctcgggca atgatggagg ccttctttct ggatcagaaa aagaaactat gcaaaatctt 420
aatgatagat tagcttccta cctggataag gtgcgagctc tagaagaggc taatactgag 480
ctagaaaata aaattcgaga atggtatgaa acacgaggaa ctgggactgc agatgcttca 540
cagagcgatt acagcaaata ttatccactg attgaagacc tcaggaataa gatcatttca 600
gccagcattg gaaatgccca gctcctcttg cagattgaca atgcgagact agctgctgag 660
gacttcagga tgaagtatga gaatgaactg gccctgcgcc agggcgtaga ggccgacatc 720
aatggcctgc gccgggtgct ggacgagctg accctgacca ggaccgacct ggagatgcag 780
atcgagagcc tgaacgagga gctggcctac atgaagaaga accacgagga tgagctccaa 840
agcttccggg tgggcggccc aggcgaggtc agcgtagaaa tggacgctgc cccccggagtg 900
gacctcacca ggctcctcaa tgatatgcgg gcgcagtatg aaaccatcgc tgagcagaat 960
cggaaggacg ctgaagcctg gttcattgaa aagagcgggg agctccgtaa ggagattagc 1020
accaacaccg agcagcttca gtccagcaag agcgaggtca ccgacctgcg tcgcgccttt 1080
cagaacctgg agatcgagct acagtcccag ctcgccatga agaaatccct ggaggactcc 1140
ttggccgaag ccgagggcga ttactgcgcg cagctgtccc aggtgcagca gctcatcagc 1200
aacctggagg cacagctgct ccaggtgcgc gcggacgcag agcgccagaa cgtggaccac 1260
cagcggctgc tgaatgtcaa ggcccgcctg gagctggaga ttgagaccta ccgccgcctg 1320
ctggacgggg aggcccaagg tgatggtttg gaggaaagtt tatttgtgac agactccaaa 1380
tcacaagcac agtcaactga ttcctctaaa gacccaacca aaacccgaaa aatcaagaca 1440
gttgtgcagg agatggtgaa tggtgaggtg gtctcatctc aagttcagga aattgaagaa 1500
ctaatgtaaa atttcacaag atctgcccca tgattggttc cttaggaaca agaaatttac 1560
aagtagaaat tattcctttc agagtaacat gctgtattac ttcaatccct atttttgtct 1620
gttccatttt ctttggattc cctattcaca ttgaatcctt tttgcccttc tgaaacaata 1680
ttcagtcaca agtcatttg gtcatgttgg tctttgtaac aaatcaaaat taccttatat 1740
ccttctggac aactggagta gtcttttaac gaactttctt ctggtaaccc ggaatatttt 1800
cttaatcata gagctttact caagtagtat tgttttaata gagttaattg taataaaaga 1860
tgaatggtaa a 1871

<210> 26

<211> 1447

<212> DNA

<213> Homo sapiens

<400> 26
ctgcaactgg ttctgcgagg gctccttcaa tggcagcgag aaggagacta tgcagttcct 60
gaacgaccgc ctggccagct acctggagaa ggtgcgtcac gtggagcggg acaacgcgga 120
gctggagaac ctcatccggg agcggtctca gcagcaggag cccttgctgt gccccagcta 180
ccagtcctac ttcaagacca ttgaggagct ccagcagaag atcctgtgca gcaagtctga 240
gaatgccagg ctggtggtgc agatcgacaa tgccaagctg gctgcagatg acttcagaac 300
caagtaccag acggagcagt ccctgcggca gctggtggag tccgacatca acagcctgcg 360
caggattctg gatgagctga ccctgtgcag gtctgacctg gaggcccaga tggagtccct 420

gaaggaggag ctgctgtccc tcaagcagaa ccatgagcag gaagtcaaca ccttgcgctg 480
ccagcttgga gaccgcctca acgtggaggt ggacgctgct cccgctgtgg acctgaacca 540
ggtcctgaac gagaccagga atcagtatga ggccctggtg gaaaccaacc gcagggaagt 600
ggagcaatgg ttcgccacgc agaccgagga gctgaacaag caggtggtat ccagctcgga 660
gcagctgcag tcctaccagg cggagatcat cgagctgaga cgcacagtca atgccctgga 720
gatcgagctg caggcccagc acaacctgcg atactctctg gaaaacacgc tgacagagag 780
cgaggcccgc tacagctccc agctgtccca ggtgcagagc ctgatcacca acgtggagtc 840
ccagctggcg gagatccgca gtgacctgga gcggcagaac caggagtatc aggtgctgct 900
ggacgtgcgg gcgcggctgg agtgtgagat caacacatac cggagcctgc tggagagcga 960
ggactgcaag ctgccctcca acccctgcgc caccaccaat gcatgtgaaa agcccattgg 1020
atcctgtgtc accaatcctt gtggtcctcg ttcccgctgt gggccttgca acacctttgg 1080
gtactagata ccctggggcc agcagaagta tagcatgaag acagaactac catcggtggg 1140
ccagttctgc ctctctgaca accatcagcc accggacccc accccgaggc atcaccacaa 1200
atcatggtct ggaaggagaa caaatgccca gcgtttgggt ctgactctga gcctagggct 1260
actgatcctc ctcaccccag gtccctctcc tgtagtcagt ctgagttctg atggtcagag 1320
gttggagctg tgacagtggc atacgaggtg ttttgttctc tctgctgctt ctacctttat 1380
tgcagttccc caaatcgcct aataaacttt cctcttgcaa agcagacaaa aaaaaaaaaa 1440
aaaaaaa 1447

<210> 27

<211> 261

<212> PRT

<213> Homo sapiens

<400> 27

Met Asn Pro Asn Cys Ala Arg Cys Gly Lys Ile Val Tyr Pro Thr Glu
1 5 10 15
Lys Val Asn Cys Leu Asp Lys Phe Trp His Lys Ala Cys Phe His Cys
20 25 30
Glu Thr Cys Lys Met Thr Leu Asn Met Lys Asn Tyr Lys Gly Tyr Glu
35 40 45
Lys Lys Pro Tyr Cys Asn Ala His Tyr Pro Lys Gln Ser Phe Thr Met
50 55 60
Val Ala Asp Thr Pro Glu Asn Leu Arg Leu Lys Gln Gln Ser Glu Leu
65 70 75 80
Gln Ser Gln Val Arg Tyr Lys Glu Glu Phe Glu Lys Asn Lys Gly Lys
85 90 95
Gly Phe Ser Val Val Ala Asp Thr Pro Glu Leu Gln Arg Ile Lys Lys
100 105 110
Thr Gln Asp Gln Ile Ser Asn Ile Lys Tyr His Glu Glu Phe Glu Lys
115 120 125
Ser Arg Met Gly Pro Ser Gly Gly Glu Gly Met Glu Pro Glu Arg Arg
130 135 140
Asp Ser Gln Asp Gly Ser Ser Tyr Arg Arg Pro Leu Glu Gln Gln Gln
145 150 155 160
Pro His His Ile Pro Thr Ser Ala Pro Val Tyr Gln Gln Pro Gln Gln
165 170 175
Gln Pro Val Ala Gln Ser Tyr Gly Gly Tyr Lys Glu Pro Ala Ala Pro
180 185 190
Val Ser Ile Gln Arg Ser Ala Pro Gly Gly Gly Gly Lys Arg Tyr Arg
195 200 205
Ala Val Tyr Asp Tyr Ser Ala Ala Asp Glu Asp Glu Val Ser Phe Gln
210 215 220
Asp Gly Asp Thr Ile Val Asn Val Gln Gln Ile Asp Asp Gly Trp Met
225 230 235 240
Tyr Gly Thr Val Glu Arg Thr Gly Asp Thr Gly Met Leu Pro Ala Asn
245 250 255
Tyr Val Glu Ala Ile
260

<210> 28

<211> 478

<212> PRT

<213> Homo sapiens

<400> 28
Met Val Gln Lys Thr Ser Met Ser Arg Gly Pro Tyr Pro Pro Ser Gln
1 5 10 15
Glu Ile Pro Met Glu Val Phe Asp Pro Ser Pro Gln Gly Lys Tyr Ser
20 25 30
Lys Arg Lys Gly Arg Phe Lys Arg Ser Asp Gly Ser Thr Ser Ser Asp
35 40 45
Thr Thr Ser Asn Ser Phe Val Arg Gln Gly Ser Ala Glu Ser Tyr Thr
50 55 60
Ser Arg Pro Ser Asp Ser Asp Val Ser Leu Glu Glu Asp Arg Glu Ala
65 70 75 80
Leu Arg Lys Glu Ala Glu Arg Gln Ala Leu Ala Gln Leu Glu Lys Ala
85 90 95
Lys Thr Lys Pro Val Ala Phe Ala Val Arg Thr Asn Val Gly Tyr Asn
100 105 110
Pro Ser Pro Gly Asp Glu Val Pro Val Gln Gly Val Ala Ile Thr Phe
115 120 125
Glu Pro Lys Asp Phe Leu His Ile Lys Glu Lys Tyr Asn Asn Asp Trp
130 135 140
Trp Ile Gly Arg Leu Val Lys Glu Gly Cys Glu Val Gly Phe Ile Pro
145 150 155 160
Ser Pro Val Lys Leu Asp Ser Leu Arg Leu Leu Gln Glu Gln Lys Leu
165 170 175
Arg Gln Asn Arg Leu Gly Ser Ser Lys Ser Gly Asp Asn Ser Ser Ser
180 185 190
Ser Leu Gly Asp Val Val Thr Gly Thr Arg Arg Pro Thr Pro Pro Ala
195 200 205
Ser Ala Lys Gln Lys Gln Lys Ser Thr Glu His Val Pro Pro Tyr Asp
210 215 220
Val Val Pro Ser Met Arg Pro Ile Ile Leu Val Gly Pro Ser Leu Lys
225 230 235 240
Gly Tyr Glu Val Thr Asp Met Met Gln Lys Ala Leu Phe Asp Phe Leu
245 250 255
Lys His Arg Phe Asp Gly Arg Ile Ser Ile Thr Arg Val Thr Ala Asp
260 265 270
Ile Ser Leu Ala Lys Arg Ser Val Leu Asn Asn Pro Ser Lys His Ile
275 280 285
Ile Ile Glu Arg Ser Asn Thr Arg Ser Ser Leu Ala Glu Val Gln Ser
290 295 300
Glu Ile Glu Arg Ile Phe Glu Leu Ala Arg Thr Leu Gln Leu Val Ala
305 310 315 320
Leu Asp Ala Asp Thr Ile Asn His Pro Ala Gln Leu Ser Lys Thr Ser
325 330 335
Leu Ala Pro Ile Ile Val Tyr Ile Lys Ile Thr Ser Pro Lys Val Leu
340 345 350
Gln Arg Leu Ile Lys Ser Arg Gly Lys Ser Gln Ser Lys His Leu Asn
355 360 365
Val Gln Ile Ala Ala Ser Glu Lys Leu Ala Gln Cys Pro Pro Glu Met
370 375 380
Phe Asp Ile Ile Leu Asp Glu Asn Gln Leu Glu Asp Ala Cys Glu His
385 390 395 400
Leu Ala Glu Tyr Leu Glu Ala Tyr Trp Lys Ala Thr His Pro Pro Ser
405 410 415
Ser Thr Pro Pro Asn Pro Leu Leu Asn Arg Thr Met Ala Thr Ala Ala
420 425 430
Leu Arg Arg Ser Pro Ala Pro Val Ser Asn Leu Gln Val Gln Val Leu
435 440 445
Thr Ser Leu Arg Arg Asn Leu Gly Phe Trp Gly Gly Leu Glu Ser Ser
450 455 460

Gln Arg Gly Ser Val Val Pro Gln Glu Gln Glu His Ala Met
465 470 475

<210> 29

<211> 196

<212> PRT

<213> Homo sapiens

<400> 29
Met Ser Met Leu Arg Leu Gln Lys Arg Leu Ala Ser Ser Val Leu Arg
1 5 10 15
Cys Gly Lys Lys Lys Val Trp Leu Asp Pro Asn Glu Thr Asn Glu Ile
20 25 30
Ala Asn Ala Asn Ser Arg Gln Gln Ile Arg Lys Leu Ile Lys Asp Gly
35 40 45
Leu Ile Ile Arg Lys Pro Val Thr Val His Ser Arg Ala Arg Cys Arg
50 55 60
Lys Asn Thr Leu Ala Arg Arg Lys Gly Arg His Met Gly Ile Gly Lys
65 70 75 80
Arg Lys Gly Thr Ala Asn Ala Arg Met Pro Glu Lys Val Thr Trp Met
85 90 95
Arg Arg Met Arg Ile Leu Arg Arg Leu Leu Arg Arg Tyr Arg Glu Ser
100 105 110
Lys Lys Ile Asp Arg His Met Tyr His Ser Leu Tyr Leu Lys Val Lys
115 120 125
Gly Asn Val Phe Lys Asn Lys Arg Ile Leu Met Glu His Ile His Lys
130 135 140
Leu Lys Ala Asp Lys Ala Arg Lys Lys Leu Leu Ala Asp Gln Ala Glu
145 150 155 160
Ala Arg Arg Ser Lys Thr Lys Glu Ala Arg Lys Arg Arg Glu Glu Arg
165 170 175
Leu Gln Ala Lys Lys Glu Glu Ile Ile Lys Thr Leu Ser Lys Glu Glu
180 185 190
Glu Thr Lys Lys
195
<210> 30

<211> 1566

<212> PRT

<213> Homo sapiens

<400> 30
Met Ser Ser Leu Leu Glu Arg Leu His Ala Lys Phe Asn Gln Asn Arg
1 5 10 15
Pro Trp Ser Glu Thr Ile Lys Leu Val Arg Gln Val Met Glu Lys Arg
20 25 30
Val Val Met Ser Ser Gly Gly His Gln His Leu Val Ser Cys Leu Glu
35 40 45
Thr Leu Gln Lys Ala Leu Lys Val Thr Ser Leu Pro Ala Met Thr Asp
50 55 60
Arg Leu Glu Ser Ile Ala Gly Gln Asn Gly Leu Gly Ser His Leu Ser
65 70 75 80
Ala Ser Gly Thr Glu Cys Tyr Ile Thr Ser Asp Met Phe Tyr Val Glu
85 90 95
Val Gln Leu Asp Pro Ala Gly Gln Leu Cys Asp Val Lys Val Ala His
100 105 110
His Gly Glu Asn Pro Val Ser Cys Pro Glu Leu Val Gln Gln Leu Arg
115 120 125

Glu Lys Asn Ser Asp Glu Phe Ser Lys His Leu Lys Gly Leu Val Asn
130 135 140
Leu Tyr Asn Leu Pro Gly Asp Asn Lys Leu Lys Thr Lys Met Tyr Leu
145 150 155 160
Ala Leu Gln Ser Leu Glu Gln Asp Leu Ser Lys Met Ala Ile Met Tyr
165 170 175
Trp Lys Ala Thr Asn Ala Gly Pro Leu Asp Lys Ile Leu His Gly Ser
180 185 190
Val Gly Tyr Leu Thr Pro Arg Ser Gly Gly His Leu Met Asn Leu Lys
195 200 205
Tyr Tyr Val Ser Pro Ser Asp Leu Leu Asp Asp Lys Thr Ala Ser Pro
210 215 220
Ile Ile Leu His Glu Asn Asn Val Ser Arg Ser Leu Gly Met Asn Ala
225 230 235 240
Ser Val Thr Ile Glu Gly Thr Ser Ala Val Tyr Lys Leu Pro Ile Ala
245 250 255
Pro Leu Ile Met Gly Ser His Pro Val Asp Asn Lys Trp Thr Pro Ser
260 265 270
Phe Ser Ser Ile Thr Ser Ala Asn Ser Val Asp Leu Pro Ala Cys Phe
275 280 285
Phe Leu Lys Phe Pro Gln Pro Ile Pro Val Ser Arg Ala Phe Val Gln
290 295 300
Lys Leu Gln Asn Cys Thr Gly Ile Pro Leu Phe Glu Thr Gln Pro Thr
305 310 315 320
Tyr Ala Pro Leu Tyr Glu Leu Ile Thr Gln Phe Glu Leu Ser Lys Asp
325 330 335
Pro Asp Pro Ile Pro Leu Asn His Asn Met Arg Phe Tyr Ala Ala Leu
340 345 350
Pro Gly Gln Gln His Cys Tyr Phe Leu Asn Lys Asp Ala Pro Leu Pro
355 360 365
Asp Gly Arg Ser Leu Gln Gly Thr Leu Val Ser Lys Ile Thr Phe Gln
370 375 380
His Pro Gly Arg Val Pro Leu Ile Leu Asn Leu Ile Arg His Gln Val
385 390 395 400
Ala Tyr Asn Thr Leu Ile Gly Ser Cys Val Lys Arg Thr Ile Leu Lys
405 410 415
Glu Asp Ser Pro Gly Leu Leu Gln Phe Glu Val Cys Pro Leu Ser Glu
420 425 430
Ser Arg Phe Ser Val Ser Phe Gln His Pro Val Asn Asp Ser Leu Val
435 440 445
Cys Val Val Met Asp Val Gln Gly Leu Thr His Val Ser Cys Lys Leu
450 455 460
Tyr Lys Gly Leu Ser Asp Ala Leu Ile Cys Thr Asp Asp Phe Ile Ala
465 470 475 480
Lys Val Val Gln Arg Cys Met Ser Ile Pro Val Thr Met Arg Ala Ile
485 490 495
Arg Arg Lys Ala Glu Thr Ile Gln Ala Asp Thr Pro Ala Leu Ser Leu
500 505 510
Ile Ala Glu Thr Val Glu Asp Met Val Lys Lys Asn Leu Pro Pro Ala
515 520 525
Ser Ser Pro Gly Tyr Gly Met Thr Thr Gly Asn Asn Pro Met Ser Gly
530 535 540
Thr Thr Thr Ser Thr Asn Thr Phe Pro Gly Gly Pro Ile Ala Thr Leu
545 550 555 560
Phe Asn Met Ser Met Ser Ile Lys Asp Arg His Glu Ser Val Gly His
565 570 575
Gly Glu Asp Phe Ser Lys Val Ser Gln Asn Pro Ile Leu Thr Ser Leu
580 585 590
Leu Gln Ile Thr Gly Asn Gly Gly Ser Thr Ile Gly Ser Ser Pro Thr
595 600 605
Pro Pro His His Thr Pro Pro Pro Val Ser Ser Met Ala Gly Asn Thr
610 615 620
Lys Asn His Pro Met Leu Met Asn Leu Leu Lys Asp Asn Pro Ala Gln
625 630 635 640
Asp Phe Ser Thr Leu Tyr Gly Ser Ser Pro Leu Glu Arg Gln Asn Ser
645 650 655

Ser Ser Gly Ser Pro Arg Met Glu Ile Cys Ser Gly Ser Asn Lys Thr
660 665 670
Lys Lys Lys Lys Ser Ser Arg Leu Pro Pro Glu Lys Pro Lys His Gln
675 680 685
Thr Glu Asp Asp Phe Gln Arg Glu Leu Phe Ser Met Asp Val Asp Ser
690 695 700
Gln Asn Pro Ile Phe Asp Val Asn Met Thr Ala Asp Thr Leu Asp Thr
705 710 715 720
Pro His Ile Thr Pro Ala Pro Ser Gln Cys Ser Thr Pro Pro Thr Thr
725 730 735
Tyr Pro Gln Pro Val Pro His Pro Gln Pro Ser Ile Gln Arg Met Val
740 745 750
Arg Leu Ser Ser Ser Asp Ser Ile Gly Pro Asp Val Thr Asp Ile Leu
755 760 765
Ser Asp Ile Ala Glu Glu Ala Ser Lys Leu Pro Ser Thr Ser Asp Asp
770 775 780
Cys Pro Ala Ile Gly Thr Pro Leu Arg Asp Ser Ser Ser Ser Gly His
785 790 795 800
Ser Gln Ser Thr Leu Phe Asp Ser Asp Val Phe Gln Thr Asn Asn Asn
805 810 815
Glu Asn Pro Tyr Thr Asp Pro Ala Asp Leu Ile Ala Asp Ala Ala Gly
820 825 830
Ser Pro Ser Ser Asp Ser Pro Thr Asn His Phe Phe His Asp Gly Val
835 840 845
Asp Phe Asn Pro Asp Leu Leu Asn Ser Gln Ser Gln Ser Gly Phe Gly
850 855 860
Glu Glu Tyr Phe Asp Glu Ser Ser Gln Ser Gly Asp Asn Asp Asp Phe
865 870 875 880
Lys Gly Phe Ala Ser Gln Ala Leu Asn Thr Leu Gly Val Pro Met Leu
885 890 895
Gly Gly Asp Asn Gly Glu Thr Lys Phe Lys Gly Asn Asn Gln Ala Asp
900 905 910
Thr Val Asp Phe Ser Ile Ile Ser Val Ala Gly Lys Ala Leu Ala Pro
915 920 925
Ala Asp Leu Met Glu His His Ser Gly Ser Gln Gly Pro Leu Leu Thr
930 935 940
Thr Gly Asp Leu Gly Lys Glu Lys Thr Gln Lys Arg Val Lys Glu Gly
945 950 955 960
Asn Gly Thr Ser Asn Ser Thr Leu Ser Gly Pro Gly Leu Asp Ser Lys
965 970 975
Pro Gly Lys Arg Ser Arg Thr Pro Ser Asn Asp Gly Lys Ser Lys Asp
980 985 990
Lys Pro Pro Lys Arg Lys Lys Ala Asp Thr Glu Gly Lys Ser Pro Ser
995 1000 1005
His Ser Ser Ser Asn Arg Pro Phe Thr Pro Pro Thr Ser Thr Gly
1010 1015 1020
Gly Ser Lys Ser Pro Gly Ser Ala Gly Arg Ser Gln Thr Pro Pro
1025 1030 1035
Gly Val Ala Thr Pro Pro Ile Pro Lys Ile Thr Ile Gln Ile Pro
1040 1045 1050
Lys Gly Thr Val Met Val Gly Lys Pro Ser Ser His Ser Gln Tyr
1055 1060 1065
Thr Ser Ser Gly Ser Val Ser Ser Ser Gly Ser Lys Ser His His
1070 1075 1080
Ser His Ser Ser Ser Ser Ser Ser Ser Ala Ser Thr Ser Gly Lys
1085 1090 1095
Met Lys Ser Ser Lys Ser Glu Gly Ser Ser Ser Ser Lys Leu Ser
1100 1105 1110
Ser Ser Met Tyr Ser Ser Gln Gly Ser Ser Gly Ser Ser Gln Ser
1115 1120 1125
Lys Asn Ser Ser Gln Ser Gly Gly Lys Pro Gly Ser Ser Pro Ile
1130 1135 1140
Thr Lys His Gly Leu Ser Ser Gly Ser Ser Ser Thr Lys Met Lys
1145 1150 1155
Pro Gln Gly Lys Pro Ser Ser Leu Met Asn Pro Ser Leu Ser Lys
1160 1165 1170

Pro Asn Ile Ser Pro Ser His Ser Arg Pro Pro Gly Gly Ser Asp
1175 1180 1185
Lys Leu Ala Ser Pro Met Lys Pro Val Pro Gly Thr Pro Pro Ser
1190 1195 1200
Ser Lys Ala Lys Ser Pro Ile Ser Ser Gly Ser Gly Gly Ser His
1205 1210 1215
Met Ser Gly Thr Ser Ser Ser Ser Gly Met Lys Ser Ser Ser Gly
1220 1225 1230
Leu Gly Ser Ser Gly Ser Leu Ser Gln Lys Thr Pro Pro Ser Ser
1235 1240 1245
Asn Ser Cys Thr Ala Ser Ser Ser Ser Phe Ser Ser Ser Gly Ser
1250 1255 1260
Ser Met Ser Ser Ser Gln Asn Gln His Gly Ser Ser Lys Gly Lys
1265 1270 1275
Ser Pro Ser Arg Asn Lys Lys Pro Ser Leu Thr Ala Val Ile Asp
1280 1285 1290
Lys Leu Lys His Gly Val Val Thr Ser Gly Pro Gly Gly Glu Asp
1295 1300 1305
Pro Leu Asp Gly Gln Met Gly Val Ser Thr Asn Ser Ser Ser His
1310 1315 1320
Pro Met Ser Ser Lys His Asn Met Ser Gly Gly Glu Phe Gln Gly
1325 1330 1335
Lys Arg Glu Lys Ser Asp Lys Asp Lys Ser Lys Val Ser Thr Ser
1340 1345 1350
Gly Ser Ser Val Asp Ser Ser Lys Lys Thr Ser Glu Ser Lys Asn
1355 1360 1365
Val Gly Ser Thr Gly Val Ala Lys Ile Ile Ile Ser Lys His Asp
1370 1375 1380
Gly Gly Ser Pro Ser Ile Lys Ala Lys Val Thr Leu Gln Lys Pro
1385 1390 1395
Gly Glu Ser Ser Gly Glu Gly Leu Arg Pro Gln Met Ala Ser Ser
1400 1405 1410
Lys Asn Tyr Gly Ser Pro Leu Ile Ser Gly Ser Thr Pro Lys His
1415 1420 1425
Glu Arg Gly Ser Pro Ser His Ser Lys Ser Pro Ala Tyr Thr Pro
1430 1435 1440
Gln Asn Leu Asp Ser Glu Ser Glu Ser Gly Ser Ser Ile Ala Glu
1445 1450 1455
Lys Ser Tyr Gln Asn Ser Pro Ser Ser Asp Asp Gly Ile Arg Pro
1460 1465 1470
Leu Pro Glu Tyr Ser Thr Glu Lys His Lys Lys His Lys Lys Glu
1475 1480 1485
Lys Lys Lys Val Lys Asp Lys Asp Arg Asp Arg Asp Arg Asp Lys
1490 1495 1500
Asp Arg Asp Lys Lys Lys Ser His Ser Ile Lys Pro Glu Ser Trp
1505 1510 1515
Ser Lys Ser Pro Ile Ser Ser Asp Gln Ser Leu Ser Met Thr Ser
1520 1525 1530
Asn Thr Ile Leu Ser Ala Asp Arg Pro Ser Arg Leu Ser Pro Asp
1535 1540 1545
Phe Met Ile Gly Glu Glu Asp Asp Asp Leu Met Asp Val Ala Leu
1550 1555 1560
Ile Gly Asn
1565

<210> 31

<211> 1490

<212> PRT

<213> Homo sapiens

<400> 31

```
Met Pro Asn Ser Glu Arg His Gly Gly Lys Lys Asp Gly Ser Gly Gly
1               5                   10                  15
Ala Ser Gly Thr Leu Gln Pro Ser Ser Gly Gly Gly Ser Ser Asn Ser
            20                  25                  30
Arg Glu Arg His Arg Leu Val Ser Lys His Lys Arg His Lys Ser Lys
        35                  40                  45
His Ser Lys Asp Met Gly Leu Val Thr Pro Glu Ala Ala Ser Leu Gly
    50                  55                  60
Thr Val Ile Lys Pro Leu Val Glu Tyr Asp Asp Ile Ser Ser Asp Ser
65                  70                  75                  80
Asp Thr Phe Ser Asp Asp Met Ala Phe Lys Leu Asp Arg Arg Glu Asn
                85                  90                  95
Asp Glu Arg Arg Gly Ser Asp Arg Ser Asp Arg Leu His Lys His Arg
            100                 105                 110
His His Gln His Arg Arg Ser Arg Asp Leu Leu Lys Ala Lys Gln Thr
        115                 120                 125
Glu Lys Glu Lys Ser Gln Glu Val Ser Ser Lys Ser Gly Ser Met Lys
    130                 135                 140
Asp Arg Ile Ser Gly Ser Ser Lys Arg Ser Asn Glu Glu Thr Asp Asp
145                 150                 155                 160
Tyr Gly Lys Ala Gln Val Ala Lys Ser Ser Ser Lys Glu Ser Arg Ser
                165                 170                 175
Ser Lys Leu His Lys Glu Lys Thr Arg Lys Glu Arg Glu Leu Lys Ser
            180                 185                 190
Gly His Lys Asp Arg Ser Lys Ser His Arg Lys Arg Glu Thr Pro Lys
        195                 200                 205
Ser Tyr Lys Thr Val Asp Ser Pro Lys Arg Arg Ser Arg Ser Pro His
    210                 215                 220
Arg Lys Trp Ser Asp Ser Ser Lys Gln Asp Asp Ser Pro Ser Gly Ala
225                 230                 235                 240
Ser Tyr Gly Gln Asp Tyr Asp Leu Ser Pro Ser Arg Ser His Thr Ser
                245                 250                 255
Ser Asn Tyr Asp Ser Tyr Lys Lys Ser Pro Gly Ser Thr Ser Arg Arg
            260                 265                 270
Gln Ser Val Ser Pro Pro Tyr Lys Glu Pro Ser Ala Tyr Gln Ser Ser
        275                 280                 285
Thr Arg Ser Pro Ser Pro Tyr Ser Arg Arg Gln Arg Ser Val Ser Pro
    290                 295                 300
Tyr Ser Arg Arg Arg Ser Ser Ser Tyr Glu Arg Ser Gly Ser Tyr Ser
305                 310                 315                 320
Gly Arg Ser Pro Ser Pro Tyr Gly Arg Arg Arg Ser Ser Ser Pro Phe
                325                 330                 335
Leu Ser Lys Arg Ser Leu Ser Arg Ser Pro Leu Pro Ser Arg Lys Ser
            340                 345                 350
Met Lys Ser Arg Ser Arg Ser Pro Ala Tyr Ser Arg His Ser Ser Ser
        355                 360                 365
His Ser Lys Lys Lys Arg Ser Ser Ser Arg Ser Arg His Ser Ser Ile
    370                 375                 380
Ser Pro Val Arg Leu Pro Leu Asn Ser Ser Leu Gly Ala Glu Leu Ser
385                 390                 395                 400
Arg Lys Lys Lys Glu Arg Ala Ala Ala Ala Ala Ala Ala Lys Met Asp
                405                 410                 415
Gly Lys Glu Ser Lys Gly Ser Pro Val Phe Leu Pro Arg Lys Glu Asn
            420                 425                 430
Ser Ser Val Glu Ala Lys Asp Ser Gly Leu Glu Ser Lys Lys Leu Pro
        435                 440                 445
Arg Ser Val Lys Leu Glu Lys Ser Ala Pro Asp Thr Glu Leu Val Asn
    450                 455                 460
Val Thr His Leu Asn Thr Glu Val Lys Asn Ser Ser Asp Thr Gly Lys
465                 470                 475                 480
Val Lys Leu Asp Glu Asn Ser Glu Lys His Leu Val Lys Asp Leu Lys
                485                 490                 495
Ala Gln Gly Thr Arg Asp Ser Lys Pro Ile Ala Leu Lys Glu Glu Ile
            500                 505                 510
```

Val Thr Pro Lys Glu Thr Glu Thr Ser Glu Lys Glu Thr Pro Pro Pro
515 520 525
Leu Pro Thr Ile Ala Ser Pro Pro Pro Pro Leu Pro Thr Thr Thr Pro
530 535 540
Pro Pro Gln Thr Pro Pro Leu Pro Pro Leu Pro Pro Ile Pro Ala Leu
545 550 555 560
Pro Gln Gln Pro Pro Leu Pro Pro Ser Gln Pro Ala Phe Ser Gln Val
565 570 575
Pro Ala Ser Ser Thr Ser Thr Leu Pro Pro Ser Thr His Ser Lys Thr
580 585 590
Ser Ala Val Ser Ser Gln Ala Asn Ser Gln Pro Pro Val Gln Val Ser
595 600 605
Val Lys Thr Gln Val Ser Val Thr Ala Ala Ile Pro His Leu Lys Thr
610 615 620
Ser Thr Leu Pro Pro Leu Pro Leu Pro Pro Leu Leu Pro Gly Gly Asp
625 630 635 640
Asp Met Asp Ser Pro Lys Glu Thr Leu Pro Ser Lys Pro Val Lys Lys
645 650 655
Glu Lys Glu Gln Arg Thr Arg His Leu Leu Thr Asp Leu Pro Leu Pro
660 665 670
Pro Glu Leu Pro Gly Gly Asp Leu Ser Pro Pro Asp Ser Pro Glu Pro
675 680 685
Lys Ala Ile Thr Pro Pro Gln Gln Pro Tyr Lys Lys Arg Pro Lys Ile
690 695 700
Cys Cys Pro Arg Tyr Gly Glu Arg Arg Gln Thr Glu Ser Asp Trp Gly
705 710 715 720
Lys Arg Cys Val Asp Lys Phe Asp Ile Ile Gly Ile Ile Gly Glu Gly
725 730 735
Thr Tyr Gly Gln Val Tyr Lys Ala Arg Asp Lys Asp Thr Gly Glu Leu
740 745 750
Val Ala Leu Lys Lys Val Arg Leu Asp Asn Glu Lys Glu Gly Phe Pro
755 760 765
Ile Thr Ala Ile Arg Glu Ile Lys Ile Leu Arg Gln Leu Ile His Arg
770 775 780
Ser Val Val Asn Met Lys Glu Ile Val Thr Asp Lys Gln Asp Ala Leu
785 790 795 800
Asp Phe Lys Lys Asp Lys Gly Ala Phe Tyr Leu Val Phe Glu Tyr Met
805 810 815
Asp His Asp Leu Met Gly Leu Leu Glu Ser Gly Leu Val His Phe Ser
820 825 830
Glu Asp His Ile Lys Ser Phe Met Lys Gln Leu Met Glu Gly Leu Glu
835 840 845
Tyr Cys His Lys Lys Asn Phe Leu His Arg Asp Ile Lys Cys Ser Asn
850 855 860
Ile Leu Leu Asn Asn Ser Gly Gln Ile Lys Leu Ala Asp Phe Gly Leu
865 870 875 880
Ala Arg Leu Tyr Asn Ser Glu Glu Ser Arg Pro Tyr Thr Asn Lys Val
885 890 895
Ile Thr Leu Trp Tyr Arg Pro Pro Glu Leu Leu Leu Gly Glu Glu Arg
900 905 910
Tyr Thr Pro Ala Ile Asp Val Trp Ser Cys Gly Cys Ile Leu Gly Glu
915 920 925
Leu Phe Thr Lys Lys Pro Ile Phe Gln Ala Asn Leu Glu Leu Ala Gln
930 935 940
Leu Glu Leu Ile Ser Arg Leu Cys Gly Ser Pro Cys Pro Ala Val Trp
945 950 955 960
Pro Asp Val Ile Lys Leu Pro Tyr Phe Asn Thr Met Lys Pro Lys Lys
965 970 975
Gln Tyr Arg Arg Arg Leu Arg Glu Glu Phe Ser Phe Ile Pro Ser Ala
980 985 990
Ala Leu Asp Leu Leu Asp His Met Leu Thr Leu Asp Pro Ser Lys Arg
995 1000 1005
Cys Thr Ala Glu Gln Thr Leu Gln Ser Asp Phe Leu Lys Asp Val
1010 1015 1020
Glu Leu Ser Lys Met Ala Pro Pro Asp Leu Pro His Trp Gln Asp
1025 1030 1035

Cys His Glu Leu Trp Ser Lys Lys Arg Arg Arg Gln Arg Gln Ser
1040 1045 1050
Gly Val Val Val Glu Glu Pro Pro Pro Ser Lys Thr Ser Arg Lys
1055 1060 1065
Glu Thr Thr Ser Gly Thr Ser Thr Glu Pro Val Lys Asn Ser Ser
1070 1075 1080
Pro Ala Pro Pro Gln Pro Ala Pro Gly Lys Val Glu Ser Gly Ala
1085 1090 1095
Gly Asp Ala Ile Gly Leu Ala Asp Ile Thr Gln Gln Leu Asn Gln
1100 1105 1110
Ser Glu Leu Ala Val Leu Leu Asn Leu Leu Gln Ser Gln Thr Asp
1115 1120 1125
Leu Ser Ile Pro Gln Met Ala Gln Leu Leu Asn Ile His Ser Asn
1130 1135 1140
Pro Glu Met Gln Gln Gln Leu Glu Ala Leu Asn Gln Ser Ile Ser
1145 1150 1155
Ala Leu Thr Glu Ala Thr Ser Gln Gln Gln Asp Ser Glu Thr Met
1160 1165 1170
Ala Pro Glu Glu Ser Leu Lys Glu Ala Pro Ser Ala Pro Val Ile
1175 1180 1185
Leu Pro Ser Ala Glu Gln Met Thr Leu Glu Ala Ser Ser Thr Pro
1190 1195 1200
Ala Asp Met Gln Asn Ile Leu Ala Val Leu Leu Ser Gln Leu Met
1205 1210 1215
Lys Thr Gln Glu Pro Ala Gly Ser Leu Glu Glu Asn Asn Ser Asp
1220 1225 1230
Lys Asn Ser Gly Pro Gln Gly Pro Arg Arg Thr Pro Thr Met Pro
1235 1240 1245
Gln Glu Glu Ala Ala Ala Cys Pro Pro His Ile Leu Pro Pro Glu
1250 1255 1260
Lys Arg Pro Pro Glu Pro Pro Gly Pro Pro Pro Pro Pro Pro Pro
1265 1270 1275
Pro Pro Leu Val Glu Gly Asp Leu Ser Ser Ala Pro Gln Glu Leu
1280 1285 1290
Asn Pro Ala Val Thr Ala Ala Leu Leu Gln Leu Leu Ser Gln Pro
1295 1300 1305
Glu Ala Glu Pro Pro Gly His Leu Pro His Glu His Gln Ala Leu
1310 1315 1320
Arg Pro Met Glu Tyr Ser Thr Arg Pro Arg Pro Asn Arg Thr Tyr
1325 1330 1335
Gly Asn Thr Asp Gly Pro Glu Thr Gly Phe Ser Ala Ile Asp Thr
1340 1345 1350
Asp Glu Arg Asn Ser Gly Pro Ala Leu Thr Glu Ser Leu Val Gln
1355 1360 1365
Thr Leu Val Lys Asn Arg Thr Phe Ser Gly Ser Leu Ser His Leu
1370 1375 1380
Gly Glu Ser Ser Ser Tyr Gln Gly Thr Gly Ser Val Gln Phe Pro
1385 1390 1395
Gly Asp Gln Asp Leu Arg Phe Ala Arg Val Pro Leu Ala Leu His
1400 1405 1410
Pro Val Val Gly Gln Pro Phe Leu Lys Ala Glu Gly Ser Ser Asn
1415 1420 1425
Ser Val Val His Ala Glu Thr Lys Leu Gln Asn Tyr Gly Glu Leu
1430 1435 1440
Gly Pro Gly Thr Thr Gly Ala Ser Ser Ser Gly Ala Gly Leu His
1445 1450 1455
Trp Gly Gly Pro Thr Gln Ser Ser Ala Tyr Gly Lys Leu Tyr Arg
1460 1465 1470
Gly Pro Thr Arg Val Pro Pro Arg Gly Gly Arg Gly Arg Gly Val
1475 1480 1485
Pro Tyr
1490

<210> 32

<211> 381

<212> PRT

<213> Homo sapiens

<400> 32
Met Leu Thr Arg Leu Phe Ser Glu Pro Gly Leu Leu Ser Asp Val Pro
1 5 10 15
Lys Phe Ala Ser Trp Gly Asp Gly Glu Asp Asp Glu Pro Arg Ser Asp
20 25 30
Lys Gly Asp Ala Pro Pro Pro Pro Pro Pro Ala Pro Gly Pro Gly Ala
35 40 45
Pro Gly Pro Ala Arg Ala Ala Lys Pro Val Pro Leu Arg Gly Glu Glu
50 55 60
Gly Thr Glu Ala Thr Leu Ala Glu Val Lys Glu Glu Gly Glu Leu Gly
65 70 75 80
Gly Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Gly Leu Asp Glu Ala
85 90 95
Glu Gly Glu Arg Pro Lys Lys Arg Gly Pro Lys Lys Arg Lys Met Thr
100 105 110
Lys Ala Arg Leu Glu Arg Ser Lys Leu Arg Arg Gln Lys Ala Asn Ala
115 120 125
Arg Glu Arg Asn Arg Met His Asp Leu Asn Ala Ala Leu Asp Asn Leu
130 135 140
Arg Lys Val Val Pro Cys Tyr Ser Lys Thr Gln Lys Leu Ser Lys Ile
145 150 155 160
Glu Thr Leu Arg Leu Ala Lys Asn Tyr Ile Trp Ala Leu Ser Glu Ile
165 170 175
Leu Arg Ser Gly Lys Arg Pro Asp Leu Val Ser Tyr Val Gln Thr Leu
180 185 190
Cys Lys Gly Leu Ser Gln Pro Thr Thr Asn Leu Val Ala Gly Cys Leu
195 200 205
Gln Leu Asn Ser Arg Asn Phe Leu Thr Glu Gln Gly Ala Asp Gly Ala
210 215 220
Gly Arg Phe His Gly Ser Gly Gly Pro Phe Ala Met His Pro Tyr Pro
225 230 235 240
Tyr Pro Cys Ser Arg Leu Ala Gly Ala Gln Cys Gln Ala Ala Gly Gly
245 250 255
Leu Gly Gly Gly Ala Ala His Ala Leu Arg Thr His Gly Tyr Cys Ala
260 265 270
Ala Tyr Glu Thr Leu Tyr Ala Ala Ala Gly Gly Gly Gly Ala Ser Pro
275 280 285
Asp Tyr Asn Ser Ser Glu Tyr Glu Gly Pro Leu Ser Pro Pro Leu Cys
290 295 300
Leu Asn Gly Asn Phe Ser Leu Lys Gln Asp Ser Ser Pro Asp His Glu
305 310 315 320
Lys Ser Tyr His Tyr Ser Met His Tyr Ser Ala Leu Pro Gly Ser Arg
325 330 335
His Gly His Gly Leu Val Phe Gly Ser Ser Ala Val Arg Gly Gly Val
340 345 350
His Ser Glu Asn Leu Leu Ser Tyr Asp Met His Leu His His Asp Arg
355 360 365
Gly Pro Met Tyr Glu Glu Leu Asn Ala Phe Phe His Asn
370 375 380

<210> 33

<211> 445

<212> PRT

<213> Homo sapiens

<400> 33

```
Met Ser Lys Leu Pro Arg Glu Leu Thr Arg Asp Leu Glu Arg Ser Leu
1               5                   10                  15
Pro Ala Val Ala Ser Leu Gly Ser Ser Leu Ser His Ser Gln Ser Leu
            20                  25                  30
Ser Ser His Leu Leu Pro Pro Pro Glu Lys Arg Arg Ala Ile Ser Asp
        35                  40                  45
Val Arg Arg Thr Phe Cys Leu Phe Val Thr Phe Asp Leu Leu Phe Ile
    50                  55                  60
Ser Leu Leu Trp Ile Ile Glu Leu Asn Thr Asn Thr Gly Ile Arg Lys
65                  70                  75                  80
Asn Leu Glu Gln Glu Ile Ile Gln Tyr Asn Phe Lys Thr Ser Phe Phe
                85                  90                  95
Asp Ile Phe Val Leu Ala Phe Phe Arg Phe Ser Gly Leu Leu Leu Gly
            100                 105                 110
Tyr Ala Val Leu Gln Leu Arg His Trp Trp Val Ile Ala Val Thr Thr
        115                 120                 125
Leu Val Ser Ser Ala Phe Leu Ile Val Lys Val Ile Leu Ser Glu Leu
    130                 135                 140
Leu Ser Lys Gly Ala Phe Gly Tyr Leu Leu Pro Ile Val Ser Phe Val
145                 150                 155                 160
Leu Ala Trp Leu Glu Thr Trp Phe Leu Asp Phe Lys Val Leu Pro Gln
                165                 170                 175
Glu Ala Glu Glu Glu Arg Trp Tyr Leu Ala Ala Gln Val Ala Val Ala
            180                 185                 190
Arg Gly Pro Leu Leu Phe Ser Gly Ala Leu Ser Glu Gly Gln Phe Tyr
        195                 200                 205
Ser Pro Pro Glu Ser Phe Ala Gly Ser Asp Asn Glu Ser Asp Glu Glu
    210                 215                 220
Val Ala Gly Lys Lys Ser Phe Ser Ala Gln Glu Arg Glu Tyr Ile Arg
225                 230                 235                 240
Gln Gly Lys Glu Ala Thr Ala Val Val Asp Gln Ile Leu Ala Gln Glu
                245                 250                 255
Glu Asn Trp Lys Phe Glu Lys Asn Asn Glu Tyr Gly Asp Thr Val Tyr
            260                 265                 270
Thr Ile Glu Val Pro Phe His Gly Lys Thr Phe Ile Leu Lys Thr Phe
        275                 280                 285
Leu Pro Cys Pro Ala Glu Leu Val Tyr Gln Glu Val Ile Leu Gln Pro
    290                 295                 300
Glu Arg Met Val Leu Trp Asn Lys Thr Val Thr Ala Cys Gln Ile Leu
305                 310                 315                 320
Gln Arg Val Glu Asp Asn Thr Leu Ile Ser Tyr Asp Val Ser Ala Gly
                325                 330                 335
Ala Ala Gly Gly Val Val Ser Pro Arg Asp Phe Val Asn Val Arg Arg
            340                 345                 350
Ile Glu Arg Arg Arg Asp Arg Tyr Leu Ser Ser Gly Ile Ala Thr Ser
        355                 360                 365
His Ser Ala Lys Pro Pro Thr His Lys Tyr Val Arg Gly Glu Asn Gly
    370                 375                 380
Pro Gly Gly Phe Ile Val Leu Lys Ser Ala Ser Asn Pro Arg Val Cys
385                 390                 395                 400
Thr Phe Val Trp Ile Leu Asn Thr Asp Leu Lys Gly Arg Leu Pro Arg
                405                 410                 415
Tyr Leu Ile His Gln Ser Leu Ala Ala Thr Met Phe Glu Phe Ala Phe
            420                 425                 430
His Leu Arg Gln Arg Ile Ser Glu Leu Gly Ala Arg Ala
        435                 440                 445
```

<210> 34

<211> 167

<212> PRT

<213> Homo sapiens

<400> 34
Met Ala Thr Ser Glu Leu Ser Cys Glu Val Ser Glu Glu Asn Cys Glu
1 5 10 15
Arg Arg Glu Ala Phe Trp Ala Glu Trp Lys Asp Leu Thr Leu Ser Thr
20 25 30
Arg Pro Glu Glu Gly Cys Ser Leu His Glu Glu Asp Thr Gln Arg His
35 40 45
Glu Thr Tyr His Gln Gln Gly Gln Cys Gln Val Leu Val Gln Arg Ser
50 55 60
Pro Trp Leu Met Met Arg Met Gly Ile Leu Gly Arg Gly Leu Gln Glu
65 70 75 80
Tyr Gln Leu Pro Tyr Gln Arg Val Leu Pro Leu Pro Ile Phe Thr Pro
85 90 95
Ala Lys Met Gly Ala Thr Lys Glu Glu Arg Glu Asp Thr Pro Ile Gln
100 105 110
Leu Gln Glu Leu Leu Ala Leu Glu Thr Ala Leu Gly Gly Gln Cys Val
115 120 125
Asp Arg Gln Glu Val Ala Glu Ile Thr Lys Gln Leu Pro Pro Val Val
130 135 140
Pro Val Ser Lys Pro Gly Ala Leu Arg Arg Ser Leu Ser Arg Ser Met
145 150 155 160
Ser Gln Glu Ala Gln Arg Gly
165

<210> 35

<211> 282

<212> PRT

<213> Homo sapiens

<400> 35
Met Ser Gly Ala Asp Arg Ser Pro Asn Ala Gly Ala Ala Pro Asp Ser
1 5 10 15
Ala Pro Gly Gln Ala Ala Val Ala Ser Ala Tyr Gln Arg Phe Glu Pro
20 25 30
Arg Ala Tyr Leu Arg Asn Asn Tyr Ala Pro Pro Arg Gly Asp Leu Cys
35 40 45
Asn Pro Asn Gly Val Gly Pro Trp Lys Leu Arg Cys Leu Ala Gln Thr
50 55 60
Phe Ala Thr Gly Glu Val Ser Gly Arg Thr Leu Ile Asp Ile Gly Ser
65 70 75 80
Gly Pro Thr Val Tyr Gln Leu Leu Ser Ala Cys Ser His Phe Glu Asp
85 90 95
Ile Thr Met Thr Asp Phe Leu Glu Val Asn Arg Gln Glu Leu Gly Arg
100 105 110
Trp Leu Gln Glu Glu Pro Gly Ala Phe Asn Trp Ser Met Tyr Ser Gln
115 120 125
His Ala Cys Leu Ile Glu Gly Lys Gly Glu Cys Trp Gln Asp Lys Glu
130 135 140
Arg Gln Leu Arg Ala Arg Val Lys Arg Val Leu Pro Ile Asp Val His
145 150 155 160
Gln Pro Gln Pro Leu Gly Ala Gly Ser Pro Ala Pro Leu Pro Ala Asp
165 170 175
Ala Leu Val Ser Ala Phe Cys Leu Glu Ala Val Ser Pro Asp Leu Ala
180 185 190
Ser Phe Gln Arg Ala Leu Asp His Ile Thr Thr Leu Leu Arg Pro Gly
195 200 205
Gly His Leu Leu Leu Ile Gly Ala Leu Glu Glu Ser Trp Tyr Leu Ala
210 215 220
Gly Glu Ala Arg Leu Thr Val Val Pro Val Ser Glu Glu Glu Val Arg
225 230 235 240

Glu Ala Leu Val Arg Ser Gly Tyr Lys Val Arg Asp Leu Arg Thr Tyr
245 250 255
Ile Met Pro Ala His Leu Gln Thr Gly Val Asp Asp Val Lys Gly Val
260 265 270
Phe Phe Ala Trp Ala Gln Lys Val Gly Leu
275 280

<210> 36

<211> 1255

<212> PRT

<213> Homo sapiens

<400> 36
Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1 5 10 15
Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
20 25 30
Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
35 40 45
Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
50 55 60
Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65 70 75 80
Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
85 90 95
Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
100 105 110
Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
115 120 125
Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
130 135 140
Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145 150 155 160
Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
165 170 175
Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
180 185 190
His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
195 200 205
Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
210 215 220
Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225 230 235 240
Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
245 250 255
His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
260 265 270
Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
275 280 285
Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
290 295 300
Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305 310 315 320
Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335
Pro Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His Leu Arg Glu
340 345 350
Val Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala Gly Cys Lys
355 360 365
Lys Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe Asp Gly Asp
370 375 380

Pro Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu Gln Val Phe
385 390 395 400
Glu Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser Ala Trp Pro
405 410 415
Asp Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln Val Ile Arg
420 425 430
Gly Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu Gln Gly Leu
435 440 445
Gly Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu Gly Ser Gly
450 455 460
Leu Ala Leu Ile His His Asn Thr His Leu Cys Phe Val His Thr Val
465 470 475 480
Pro Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu Leu His Thr
485 490 495
Ala Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu Ala Cys His
500 505 510
Gln Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro Thr Gln Cys
515 520 525
Val Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val Glu Glu Cys
530 535 540
Arg Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys
545 550 555 560
Leu Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys
565 570 575
Phe Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp
580 585 590
Pro Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu
595 600 605
Ser Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln
610 615 620
Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys
625 630 635 640
Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile Val Ser
645 650 655
Ala Val Val Gly Ile Leu Leu Val Val Val Leu Gly Val Val Phe Gly
660 665 670
Ile Leu Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg
675 680 685
Arg Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly
690 695 700
Ala Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu
705 710 715 720
Arg Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys
725 730 735
Gly Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile
740 745 750
Lys Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu
755 760 765
Asp Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg
770 775 780
Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu
785 790 795 800
Met Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg
805 810 815
Leu Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly
820 825 830
Met Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala
835 840 845
Arg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe
850 855 860
Gly Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp
865 870 875 880
Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg
885 890 895
Arg Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val
900 905 910

Trp Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala
915 920 925
Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro
930 935 940
Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met
945 950 955 960
Ile Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe
965 970 975
Ser Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu
980 985 990
Asp Leu Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu
995 1000 1005
Leu Glu Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr
1010 1015 1020
Leu Val Pro Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly
1025 1030 1035
Ala Gly Gly Met Val His His Arg His Arg Ser Ser Ser Thr Arg
1040 1045 1050
Ser Gly Gly Gly Asp Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu
1055 1060 1065
Glu Ala Pro Arg Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly Ser
1070 1075 1080
Asp Val Phe Asp Gly Asp Leu Gly Met Gly Ala Ala Lys Gly Leu
1085 1090 1095
Gln Ser Leu Pro Thr His Asp Pro Ser Pro Leu Gln Arg Tyr Ser
1100 1105 1110
Glu Asp Pro Thr Val Pro Leu Pro Ser Glu Thr Asp Gly Tyr Val
1115 1120 1125
Ala Pro Leu Thr Cys Ser Pro Gln Pro Glu Tyr Val Asn Gln Pro
1130 1135 1140
Asp Val Arg Pro Gln Pro Pro Ser Pro Arg Glu Gly Pro Leu Pro
1145 1150 1155
Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu Arg Ala Lys Thr Leu
1160 1165 1170
Ser Pro Gly Lys Asn Gly Val Val Lys Asp Val Phe Ala Phe Gly
1175 1180 1185
Gly Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln Gly Gly Ala
1190 1195 1200
Ala Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala Phe Asp
1205 1210 1215
Asn Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala Pro
1220 1225 1230
Pro Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr
1235 1240 1245
Leu Gly Leu Asp Val Pro Val
1250 1255

<210> 37

<211> 532

<212> PRT

<213> Homo sapiens

<400> 37
Met Glu Leu Asp Leu Ser Pro Pro His Leu Ser Ser Ser Pro Glu Asp
1 5 10 15
Leu Trp Pro Ala Pro Gly Thr Pro Pro Gly Thr Pro Arg Pro Pro Asp
20 25 30
Thr Pro Leu Pro Glu Glu Val Lys Arg Ser Gln Pro Leu Leu Ile Pro
35 40 45
Thr Thr Gly Arg Lys Leu Arg Glu Glu Glu Arg Arg Ala Thr Ser Leu
50 55 60

Pro Ser Ile Pro Asn Pro Phe Pro Glu Leu Cys Ser Pro Pro Ser Gln
65 70 75 80
Ser Pro Ile Leu Gly Gly Pro Ser Ser Ala Arg Gly Leu Leu Pro Arg
85 90 95
Asp Ala Ser Arg Pro His Val Val Lys Val Tyr Ser Glu Asp Gly Ala
100 105 110
Cys Arg Ser Val Glu Val Ala Ala Gly Ala Thr Ala Arg His Val Cys
115 120 125
Glu Met Leu Val Gln Arg Ala His Ala Leu Ser Asp Glu Thr Trp Gly
130 135 140
Leu Val Glu Cys His Pro His Leu Ala Leu Glu Arg Gly Leu Glu Asp
145 150 155 160
His Glu Ser Val Val Glu Val Gln Ala Ala Trp Pro Val Gly Gly Asp
165 170 175
Ser Arg Phe Val Phe Arg Lys Asn Phe Ala Lys Tyr Glu Leu Phe Lys
180 185 190
Ser Ser Pro His Ser Leu Phe Pro Glu Lys Met Val Ser Ser Cys Leu
195 200 205
Asp Ala His Thr Gly Ile Ser His Glu Asp Leu Ile Gln Asn Phe Leu
210 215 220
Asn Ala Gly Ser Phe Pro Glu Ile Gln Gly Phe Leu Gln Leu Arg Gly
225 230 235 240
Ser Gly Arg Lys Leu Trp Lys Arg Phe Phe Cys Phe Leu Arg Arg Ser
245 250 255
Gly Leu Tyr Tyr Ser Thr Lys Gly Thr Ser Lys Asp Pro Arg His Leu
260 265 270
Gln Tyr Val Ala Asp Val Asn Glu Ser Asn Val Tyr Val Val Thr Gln
275 280 285
Gly Arg Lys Leu Tyr Gly Met Pro Thr Asp Phe Gly Phe Cys Val Lys
290 295 300
Pro Asn Lys Leu Arg Asn Gly His Lys Gly Leu Arg Ile Phe Cys Ser
305 310 315 320
Glu Asp Glu Gln Ser Arg Thr Cys Trp Leu Ala Ala Phe Arg Leu Phe
325 330 335
Lys Tyr Gly Val Gln Leu Tyr Lys Asn Tyr Gln Gln Ala Gln Ser Arg
340 345 350
His Leu His Pro Ser Cys Leu Gly Ser Pro Pro Leu Arg Ser Ala Ser
355 360 365
Asp Asn Thr Leu Val Ala Met Asp Phe Ser Gly His Ala Gly Arg Val
370 375 380
Ile Glu Asn Pro Arg Glu Ala Leu Ser Val Ala Leu Glu Glu Ala Gln
385 390 395 400
Ala Trp Arg Lys Lys Thr Asn His Arg Leu Ser Leu Pro Met Pro Ala
405 410 415
Ser Gly Thr Ser Leu Ser Ala Ala Ile His Arg Thr Gln Leu Trp Phe
420 425 430
His Gly Arg Ile Ser Arg Glu Glu Ser Gln Arg Leu Ile Gly Gln Gln
435 440 445
Gly Leu Val Asp Gly Leu Phe Leu Val Arg Glu Ser Gln Arg Asn Pro
450 455 460
Gln Gly Phe Val Leu Ser Leu Cys His Leu Gln Lys Val Lys His Tyr
465 470 475 480
Leu Ile Leu Pro Ser Glu Glu Glu Gly Arg Leu Tyr Phe Ser Met Asp
485 490 495
Asp Gly Gln Thr Arg Phe Thr Asp Leu Leu Gln Leu Val Glu Phe His
500 505 510
Gln Leu Asn Arg Gly Ile Leu Pro Cys Leu Leu Arg His Cys Cys Thr
515 520 525
Arg Val Ala Leu
530

<210> 38

<211> 534

<212> PRT

<213> Homo sapiens

<400> 38

```
Met Lys Gln Glu Gly Ser Ala Arg Arg Arg Gly Ala Asp Lys Ala Lys
1               5                   10                  15
Pro Pro Pro Gly Gly Gly Glu Gln Glu Pro Pro Pro Pro Pro Ala Pro
            20                  25                  30
Gln Asp Val Glu Met Lys Glu Glu Ala Ala Thr Gly Gly Gly Ser Thr
        35                  40                  45
Gly Glu Ala Asp Gly Lys Thr Ala Ala Ala Ala Val Glu His Ser Gln
    50                  55                  60
Arg Glu Leu Asp Thr Val Thr Leu Glu Asp Ile Lys Glu His Val Lys
65                  70                  75                  80
Gln Leu Glu Lys Ala Val Ser Gly Lys Glu Pro Arg Phe Val Leu Arg
                85                  90                  95
Ala Leu Arg Met Leu Pro Ser Thr Ser Arg Arg Leu Asn His Tyr Val
            100                 105                 110
Leu Tyr Lys Ala Val Gln Gly Phe Phe Thr Ser Asn Asn Ala Thr Arg
        115                 120                 125
Asp Phe Leu Leu Pro Phe Leu Glu Glu Pro Met Asp Thr Glu Ala Asp
    130                 135                 140
Leu Gln Phe Arg Pro Arg Thr Gly Lys Ala Ala Ser Thr Pro Leu Leu
145                 150                 155                 160
Pro Glu Val Glu Ala Tyr Leu Gln Leu Leu Val Val Ile Phe Met Met
                165                 170                 175
Asn Ser Lys Arg Tyr Lys Glu Ala Gln Lys Ile Ser Asp Asp Leu Met
            180                 185                 190
Gln Lys Ile Ser Thr Gln Asn Arg Arg Ala Leu Asp Leu Val Ala Ala
        195                 200                 205
Lys Cys Tyr Tyr Tyr His Ala Arg Val Tyr Glu Phe Leu Asp Lys Leu
    210                 215                 220
Asp Val Val Arg Ser Phe Leu His Ala Arg Leu Arg Thr Ala Thr Leu
225                 230                 235                 240
Arg His Asp Ala Asp Gly Gln Ala Thr Leu Leu Asn Leu Leu Leu Arg
                245                 250                 255
Asn Tyr Leu His Tyr Ser Leu Tyr Asp Gln Ala Glu Lys Leu Val Ser
            260                 265                 270
Lys Ser Val Phe Pro Glu Gln Ala Asn Asn Asn Glu Trp Ala Arg Tyr
        275                 280                 285
Leu Tyr Tyr Thr Gly Arg Ile Lys Ala Ile Gln Leu Glu Tyr Ser Glu
    290                 295                 300
Ala Arg Arg Thr Met Thr Asn Ala Leu Arg Lys Ala Pro Gln His Thr
305                 310                 315                 320
Ala Val Gly Phe Lys Gln Thr Val His Lys Leu Leu Ile Val Val Glu
                325                 330                 335
Leu Leu Leu Gly Glu Ile Pro Asp Arg Leu Gln Phe Arg Gln Pro Ser
            340                 345                 350
Leu Lys Arg Ser Leu Met Pro Tyr Phe Leu Leu Thr Gln Ala Val Arg
        355                 360                 365
Thr Gly Asn Leu Ala Lys Phe Asn Gln Val Leu Asp Gln Phe Gly Glu
    370                 375                 380
Lys Phe Gln Ala Asp Gly Thr Tyr Thr Leu Ile Ile Arg Leu Arg His
385                 390                 395                 400
Asn Val Ile Lys Thr Gly Val Arg Met Ile Ser Leu Ser Tyr Ser Arg
                405                 410                 415
Ile Ser Leu Ala Asp Ile Ala Gln Lys Leu Gln Leu Asp Ser Pro Glu
            420                 425                 430
Asp Ala Glu Phe Ile Val Ala Lys Ala Ile Arg Asp Gly Val Ile Glu
        435                 440                 445
Ala Ser Ile Asn His Glu Lys Gly Tyr Val Gln Ser Lys Glu Met Ile
    450                 455                 460
Asp Ile Tyr Ser Thr Arg Glu Pro Gln Leu Ala Phe His Gln Arg Ile
465                 470                 475                 480
```

Ser Phe Cys Leu Asp Ile His Asn Met Ser Val Lys Ala Met Arg Phe
485 490 495
Pro Pro Lys Ser Tyr Asn Lys Asp Leu Glu Ser Ala Glu Glu Arg Arg
500 505 510
Glu Arg Glu Gln Gln Asp Leu Glu Phe Ala Lys Glu Met Ala Glu Asp
515 520 525
Asp Asp Asp Ser Phe Pro
530

<210> 39

<211> 207

<212> PRT

<213> Homo sapiens

<400> 39
Met Ala Gly Pro Ala Thr Gln Ser Pro Met Lys Leu Met Ala Leu Gln
1 5 10 15
Leu Leu Leu Trp His Ser Ala Leu Trp Thr Val Gln Glu Ala Thr Pro
20 25 30
Leu Gly Pro Ala Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys Cys Leu
35 40 45
Glu Gln Val Arg Lys Ile Gln Gly Asp Gly Ala Ala Leu Gln Glu Lys
50 55 60
Leu Val Ser Glu Cys Ala Thr Tyr Lys Leu Cys His Pro Glu Glu Leu
65 70 75 80
Val Leu Leu Gly His Ser Leu Gly Ile Pro Trp Ala Pro Leu Ser Ser
85 90 95
Cys Pro Ser Gln Ala Leu Gln Leu Ala Gly Cys Leu Ser Gln Leu His
100 105 110
Ser Gly Leu Phe Leu Tyr Gln Gly Leu Leu Gln Ala Leu Glu Gly Ile
115 120 125
Ser Pro Glu Leu Gly Pro Thr Leu Asp Thr Leu Gln Leu Asp Val Ala
130 135 140
Asp Phe Ala Thr Thr Ile Trp Gln Gln Met Glu Glu Leu Gly Met Ala
145 150 155 160
Pro Ala Leu Gln Pro Thr Gln Gly Ala Met Pro Ala Phe Ala Ser Ala
165 170 175
Phe Gln Arg Arg Ala Gly Gly Val Leu Val Ala Ser His Leu Gln Ser
180 185 190
Phe Leu Glu Val Ser Tyr Arg Val Leu Arg His Leu Ala Gln Pro
195 200 205

<210> 40

<211> 989

<212> PRT

<213> Homo sapiens

<400> 40
Met Lys Val Val Asn Leu Lys Gln Ala Ile Leu Gln Ala Trp Lys Glu
1 5 10 15
Arg Trp Ser Tyr Tyr Gln Trp Ala Ile Asn Met Lys Lys Phe Phe Pro
20 25 30
Lys Gly Ala Thr Trp Asp Ile Leu Asn Leu Ala Asp Ala Leu Leu Glu
35 40 45
Gln Ala Met Ile Gly Pro Ser Pro Asn Pro Leu Ile Leu Ser Tyr Leu
50 55 60
Lys Tyr Ala Ile Ser Ser Gln Met Val Ser Tyr Ser Ser Val Leu Thr
65 70 75 80

Ala Ile Ser Lys Phe Asp Asp Phe Ser Arg Asp Leu Cys Val Gln Ala
85 90 95
Leu Leu Asp Ile Met Asp Met Phe Cys Asp Arg Leu Ser Cys His Gly
100 105 110
Lys Ala Glu Glu Cys Ile Gly Leu Cys Arg Ala Leu Leu Ser Ala Leu
115 120 125
His Trp Leu Leu Arg Cys Thr Ala Ala Ser Ala Glu Arg Leu Arg Glu
130 135 140
Gly Leu Glu Ala Gly Thr Pro Ala Ala Gly Glu Lys Gln Leu Ala Met
145 150 155 160
Cys Leu Gln Arg Leu Glu Lys Thr Leu Ser Ser Thr Lys Asn Arg Ala
165 170 175
Leu Leu His Ile Ala Lys Leu Glu Glu Ala Ser Ser Trp Thr Ala Ile
180 185 190
Glu His Ser Leu Leu Lys Leu Gly Glu Ile Leu Thr Asn Leu Ser Asn
195 200 205
Pro Gln Leu Arg Ser Gln Ala Glu Gln Cys Gly Thr Leu Ile Arg Ser
210 215 220
Ile Pro Thr Met Leu Ser Val His Ala Glu Gln Met His Lys Thr Gly
225 230 235 240
Phe Pro Thr Val His Ala Val Ile Leu Leu Glu Gly Thr Met Asn Leu
245 250 255
Thr Gly Glu Thr Gln Ser Leu Val Glu Gln Leu Thr Met Val Lys Arg
260 265 270
Met Gln His Ile Pro Thr Pro Leu Phe Val Leu Glu Ile Trp Lys Ala
275 280 285
Cys Phe Val Gly Leu Ile Glu Ser Pro Glu Gly Thr Glu Glu Leu Lys
290 295 300
Trp Thr Ala Phe Thr Phe Leu Lys Ile Pro Gln Val Leu Val Lys Leu
305 310 315 320
Lys Lys Tyr Ser His Gly Asp Lys Asp Phe Thr Glu Asp Val Asn Cys
325 330 335
Ala Phe Glu Phe Leu Leu Lys Leu Thr Pro Leu Leu Asp Lys Ala Asp
340 345 350
Gln Arg Cys Asn Cys Asp Cys Thr Asn Phe Leu Leu Gln Glu Cys Gly
355 360 365
Lys Gln Gly Leu Leu Ser Glu Ala Ser Val Asn Asn Leu Met Ala Lys
370 375 380
Arg Lys Ala Asp Arg Glu His Ala Pro Gln Gln Lys Ser Gly Glu Asn
385 390 395 400
Ala Asn Ile Gln Pro Asn Ile Gln Leu Ile Leu Arg Ala Glu Pro Thr
405 410 415
Val Thr Asn Ile Leu Lys Thr Met Asp Ala Asp His Ser Lys Ser Pro
420 425 430
Glu Gly Leu Leu Gly Val Leu Gly His Met Leu Ser Gly Lys Ser Leu
435 440 445
Asp Leu Leu Leu Ala Ala Ala Ala Ala Thr Gly Lys Leu Lys Ser Phe
450 455 460
Ala Arg Lys Phe Ile Asn Leu Asn Glu Phe Thr Thr Tyr Gly Ser Glu
465 470 475 480
Glu Ser Thr Lys Pro Ala Ser Val Arg Ala Leu Leu Phe Asp Ile Ser
485 490 495
Phe Leu Met Leu Cys His Val Ala Gln Thr Tyr Gly Ser Glu Val Ile
500 505 510
Leu Ser Glu Ser Arg Thr Gly Ala Glu Val Pro Phe Phe Glu Thr Trp
515 520 525
Met Gln Thr Cys Met Pro Glu Glu Gly Lys Ile Leu Asn Pro Asp His
530 535 540
Pro Cys Phe Arg Pro Asp Ser Thr Lys Val Glu Ser Leu Val Ala Leu
545 550 555 560
Leu Asn Asn Ser Ser Glu Met Lys Leu Val Gln Met Lys Trp His Glu
565 570 575
Ala Cys Leu Ser Ile Ser Ala Ala Ile Leu Glu Ile Leu Asn Ala Trp
580 585 590
Glu Asn Gly Val Leu Ala Phe Glu Ser Ile Gln Lys Ile Thr Asp Asn
595 600 605

Ile Lys Gly Lys Val Cys Ser Leu Ala Val Cys Ala Val Ala Trp Leu
610 615 620
Val Ala His Val Arg Met Leu Gly Leu Asp Glu Arg Glu Lys Ser Leu
625 630 635 640
Gln Met Ile Arg Gln Leu Ala Gly Pro Leu Phe Ser Glu Asn Thr Leu
645 650 655
Gln Phe Tyr Asn Glu Arg Val Val Ile Met Asn Ser Ile Leu Glu Arg
660 665 670
Met Cys Ala Asp Val Leu Gln Gln Thr Ala Thr Gln Ile Lys Phe Pro
675 680 685
Ser Thr Gly Val Asp Thr Met Pro Tyr Trp Asn Leu Leu Pro Pro Lys
690 695 700
Arg Pro Ile Lys Glu Val Leu Thr Asp Ile Phe Ala Lys Val Leu Glu
705 710 715 720
Lys Gly Trp Val Asp Ser Arg Ser Ile His Ile Phe Asp Thr Leu Leu
725 730 735
His Met Gly Gly Val Tyr Trp Phe Cys Asn Asn Leu Ile Lys Glu Leu
740 745 750
Leu Lys Glu Thr Arg Lys Glu His Thr Leu Arg Ala Val Glu Leu Leu
755 760 765
Tyr Ser Ile Phe Cys Leu Asp Met Gln Gln Val Thr Leu Val Leu Leu
770 775 780
Gly His Ile Leu Pro Gly Leu Leu Thr Asp Ser Ser Lys Trp His Ser
785 790 795 800
Leu Met Asp Pro Pro Gly Thr Ala Leu Ala Lys Leu Ala Val Trp Cys
805 810 815
Ala Leu Ser Ser Tyr Ser Ser His Lys Gly Gln Ala Ser Thr Arg Gln
820 825 830
Lys Lys Arg His Arg Glu Asp Ile Glu Asp Tyr Ile Ser Leu Phe Pro
835 840 845
Leu Asp Asp Val Gln Pro Ser Lys Leu Met Arg Leu Leu Ser Ser Asn
850 855 860
Glu Asp Asp Ala Asn Ile Leu Ser Ser Pro Thr Asp Arg Ser Met Ser
865 870 875 880
Ser Ser Leu Ser Ala Ser Gln Leu His Thr Val Asn Met Arg Asp Pro
885 890 895
Leu Asn Arg Val Leu Ala Asn Leu Phe Leu Leu Ile Ser Ser Ile Leu
900 905 910
Gly Ser Arg Thr Ala Gly Pro His Thr Gln Phe Val Gln Trp Phe Met
915 920 925
Glu Glu Cys Val Asp Cys Leu Glu Gln Gly Gly Arg Gly Ser Val Leu
930 935 940
Gln Phe Met Pro Phe Thr Thr Val Ser Glu Leu Val Lys Val Ser Ala
945 950 955 960
Met Ser Ser Pro Lys Val Val Leu Ala Ile Thr Asp Leu Ser Leu Pro
965 970 975
Leu Gly Arg Gln Val Ala Ala Lys Ala Ile Ala Ala Leu
980 985

<210> 41

<211> 490

<212> PRT

<213> Homo sapiens

<400> 41
Met Glu Gln Lys Pro Ser Lys Val Glu Cys Gly Ser Asp Pro Glu Glu
1 5 10 15
Asn Ser Ala Arg Ser Pro Asp Gly Lys Arg Lys Arg Lys Asn Gly Gln
20 25 30
Cys Ser Leu Lys Thr Ser Met Ser Gly Tyr Ile Pro Ser Tyr Leu Asp
35 40 45

Lys Asp Glu Gln Cys Val Val Cys Gly Asp Lys Ala Thr Gly Tyr His
50 55 60
Tyr Arg Cys Ile Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Thr
65 70 75 80
Ile Gln Lys Asn Leu His Pro Thr Tyr Ser Cys Lys Tyr Asp Ser Cys
85 90 95
Cys Val Ile Asp Lys Ile Thr Arg Asn Gln Cys Gln Leu Cys Arg Phe
100 105 110
Lys Lys Cys Ile Ala Val Gly Met Ala Met Asp Leu Val Leu Asp Asp
115 120 125
Ser Lys Arg Val Ala Lys Arg Lys Leu Ile Glu Gln Asn Arg Glu Arg
130 135 140
Arg Arg Lys Glu Glu Met Ile Arg Ser Leu Gln Gln Arg Pro Glu Pro
145 150 155 160
Thr Pro Glu Glu Trp Asp Leu Ile His Ile Ala Thr Glu Ala His Arg
165 170 175
Ser Thr Asn Ala Gln Gly Ser His Trp Lys Gln Arg Arg Lys Phe Leu
180 185 190
Pro Asp Asp Ile Gly Gln Ser Pro Ile Val Ser Met Pro Asp Gly Asp
195 200 205
Lys Val Asp Leu Glu Ala Phe Ser Glu Phe Thr Lys Ile Ile Thr Pro
210 215 220
Ala Ile Thr Arg Val Val Asp Phe Ala Lys Lys Leu Pro Met Phe Ser
225 230 235 240
Glu Leu Pro Cys Glu Asp Gln Ile Ile Leu Leu Lys Gly Cys Cys Met
245 250 255
Glu Ile Met Ser Leu Arg Ala Ala Val Arg Tyr Asp Pro Glu Ser Asp
260 265 270
Thr Leu Thr Leu Ser Gly Glu Met Ala Val Lys Arg Glu Gln Leu Lys
275 280 285
Asn Gly Gly Leu Gly Val Val Ser Asp Ala Ile Phe Glu Leu Gly Lys
290 295 300
Ser Leu Ser Ala Phe Asn Leu Asp Asp Thr Glu Val Ala Leu Leu Gln
305 310 315 320
Ala Val Leu Leu Met Ser Thr Asp Arg Ser Gly Leu Leu Cys Val Asp
325 330 335
Lys Ile Glu Lys Ser Gln Glu Ala Tyr Leu Leu Ala Phe Glu His Tyr
340 345 350
Val Asn His Arg Lys His Asn Ile Pro His Phe Trp Pro Lys Leu Leu
355 360 365
Met Lys Glu Arg Glu Val Gln Ser Ser Ile Leu Tyr Lys Gly Ala Ala
370 375 380
Ala Glu Gly Arg Pro Gly Gly Ser Leu Gly Val His Pro Glu Gly Gln
385 390 395 400
Gln Leu Leu Gly Met His Val Val Gln Gly Pro Gln Val Arg Gln Leu
405 410 415
Glu Gln Gln Leu Gly Glu Ala Gly Ser Leu Gln Gly Pro Val Leu Gln
420 425 430
His Gln Ser Pro Lys Ser Pro Gln Gln Arg Leu Leu Glu Leu Leu His
435 440 445
Arg Ser Gly Ile Leu His Ala Arg Ala Val Cys Gly Glu Asp Asp Ser
450 455 460
Ser Glu Ala Asp Ser Pro Ser Ser Ser Glu Glu Glu Pro Glu Val Cys
465 470 475 480
Glu Asp Leu Ala Gly Asn Ala Ala Ser Pro
485 490

<210> 42

<211> 614

<212> PRT

<213> Homo sapiens

<400> 42

Met Thr Thr Leu Asp Ser Asn Asn Asn Thr Gly Gly Val Ile Thr Tyr
1 5 10 15
Ile Gly Ser Ser Gly Ser Ser Pro Ser Arg Thr Ser Pro Glu Ser Leu
20 25 30
Tyr Ser Asp Asn Ser Asn Gly Ser Phe Gln Ser Leu Thr Gln Gly Cys
35 40 45
Pro Thr Tyr Phe Pro Pro Ser Pro Thr Gly Ser Leu Thr Gln Asp Pro
50 55 60
Ala Arg Ser Phe Gly Ser Ile Pro Pro Ser Leu Ser Asp Asp Gly Ser
65 70 75 80
Pro Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Ser Phe Tyr Asn
85 90 95
Gly Ser Pro Pro Gly Ser Leu Gln Val Ala Met Glu Asp Ser Ser Arg
100 105 110
Val Ser Pro Ser Lys Ser Thr Ser Asn Ile Thr Lys Leu Asn Gly Met
115 120 125
Val Leu Leu Cys Lys Val Cys Gly Asp Val Ala Ser Gly Phe His Tyr
130 135 140
Gly Val Leu Ala Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Ile
145 150 155 160
Gln Gln Asn Ile Gln Tyr Lys Arg Cys Leu Lys Asn Glu Asn Cys Ser
165 170 175
Ile Val Arg Ile Asn Arg Asn Arg Cys Gln Gln Cys Arg Phe Lys Lys
180 185 190
Cys Leu Ser Val Gly Met Ser Arg Asp Ala Val Arg Phe Gly Arg Ile
195 200 205
Pro Lys Arg Glu Lys Gln Arg Met Leu Ala Glu Met Gln Ser Ala Met
210 215 220
Asn Leu Ala Asn Asn Gln Leu Ser Ser Gln Cys Pro Leu Glu Thr Ser
225 230 235 240
Pro Thr Gln His Pro Thr Pro Gly Pro Met Gly Pro Ser Pro Pro Pro
245 250 255
Ala Pro Val Pro Ser Pro Leu Val Gly Phe Ser Gln Phe Pro Gln Gln
260 265 270
Leu Thr Pro Pro Arg Ser Pro Ser Pro Glu Pro Thr Val Glu Asp Val
275 280 285
Ile Ser Gln Val Ala Arg Ala His Arg Glu Ile Phe Thr Tyr Ala His
290 295 300
Asp Lys Leu Gly Ser Ser Pro Gly Asn Phe Asn Ala Asn His Ala Ser
305 310 315 320
Gly Ser Pro Pro Ala Thr Thr Pro His Arg Trp Glu Asn Gln Gly Cys
325 330 335
Pro Pro Ala Pro Asn Asp Asn Asn Thr Leu Ala Ala Gln Arg His Asn
340 345 350
Glu Ala Leu Asn Gly Leu Arg Gln Ala Pro Ser Ser Tyr Pro Pro Thr
355 360 365
Trp Pro Pro Gly Pro Ala His His Ser Cys His Gln Ser Asn Ser Asn
370 375 380
Gly His Arg Leu Cys Pro Thr His Val Tyr Ala Ala Pro Glu Gly Lys
385 390 395 400
Ala Pro Ala Asn Ser Pro Arg Gln Gly Asn Ser Lys Asn Val Leu Leu
405 410 415
Ala Cys Pro Met Asn Met Tyr Pro His Gly Arg Ser Gly Arg Thr Val
420 425 430
Gln Glu Ile Trp Glu Asp Phe Ser Met Ser Phe Thr Pro Ala Val Arg
435 440 445
Glu Val Val Glu Phe Ala Lys His Ile Pro Gly Phe Arg Asp Leu Ser
450 455 460
Gln His Asp Gln Val Thr Leu Leu Lys Ala Gly Thr Phe Glu Val Leu
465 470 475 480
Met Val Arg Phe Ala Ser Leu Phe Asn Val Lys Asp Gln Thr Val Met
485 490 495
Phe Leu Ser Arg Thr Thr Tyr Ser Leu Gln Glu Leu Gly Ala Met Gly
500 505 510

Met Gly Asp Leu Leu Ser Ala Met Phe Asp Phe Ser Glu Lys Leu Asn
515 520 525
Ser Leu Ala Leu Thr Glu Glu Glu Leu Gly Leu Phe Thr Ala Val Val
530 535 540
Leu Val Ser Ala Asp Arg Ser Gly Met Glu Asn Ser Ala Ser Val Glu
545 550 555 560
Gln Leu Gln Glu Thr Leu Leu Arg Ala Leu Arg Ala Leu Val Leu Lys
565 570 575
Asn Arg Pro Leu Glu Thr Ser Arg Phe Thr Lys Leu Leu Leu Lys Leu
580 585 590
Pro Asp Leu Arg Thr Leu Asn Asn Met His Ser Glu Lys Leu Leu Ser
595 600 605
Phe Arg Val Asp Ala Gln
610

<210> 43

<211> 703

<212> PRT

<213> Homo sapiens

<400> 43
Met Ala Asp Arg Arg Arg Gln Arg Ala Ser Gln Asp Thr Glu Asp Glu
1 5 10 15
Glu Ser Gly Ala Ser Gly Ser Asp Ser Gly Gly Ser Pro Leu Arg Gly
20 25 30
Gly Gly Ser Cys Ser Gly Ser Ala Gly Gly Gly Gly Ser Gly Ser Leu
35 40 45
Pro Ser Gln Arg Gly Gly Arg Thr Gly Ala Leu His Leu Arg Arg Val
50 55 60
Glu Ser Gly Gly Ala Lys Ser Ala Glu Glu Ser Glu Cys Glu Ser Glu
65 70 75 80
Asp Gly Ile Glu Gly Asp Ala Val Leu Ser Asp Tyr Glu Ser Ala Glu
85 90 95
Asp Ser Glu Gly Glu Glu Gly Glu Tyr Ser Glu Glu Glu Asn Ser Lys
100 105 110
Val Glu Leu Lys Ser Glu Ala Asn Asp Ala Val Asn Ser Ser Thr Lys
115 120 125
Glu Glu Lys Gly Glu Glu Lys Pro Asp Thr Lys Ser Thr Val Thr Gly
130 135 140
Glu Arg Gln Ser Gly Asp Gly Gln Glu Ser Thr Glu Pro Val Glu Asn
145 150 155 160
Lys Val Gly Lys Lys Gly Pro Lys His Leu Asp Asp Asp Glu Asp Arg
165 170 175
Lys Asn Pro Ala Tyr Ile Pro Arg Lys Gly Leu Phe Phe Glu His Asp
180 185 190
Leu Arg Gly Gln Thr Gln Glu Glu Glu Val Arg Pro Lys Gly Arg Gln
195 200 205
Arg Lys Leu Trp Lys Asp Glu Gly Arg Trp Glu His Asp Lys Phe Arg
210 215 220
Glu Asp Glu Gln Ala Pro Lys Ser Arg Gln Glu Leu Ile Ala Leu Tyr
225 230 235 240
Gly Tyr Asp Ile Arg Ser Ala His Asn Pro Asp Asp Ile Lys Pro Arg
245 250 255
Arg Ile Arg Lys Pro Arg Tyr Gly Ser Pro Pro Gln Arg Asp Pro Asn
260 265 270
Trp Asn Gly Glu Arg Leu Asn Lys Ser His Arg His Gln Gly Leu Gly
275 280 285
Gly Thr Leu Pro Pro Arg Thr Phe Ile Asn Arg Asn Ala Ala Gly Thr
290 295 300
Gly Arg Met Ser Ala Pro Arg Asn Tyr Ser Arg Ser Gly Gly Phe Lys
305 310 315 320

Glu Gly Arg Ala Gly Phe Arg Pro Val Glu Ala Gly Gly Gln His Gly
325 330 335
Gly Arg Ser Gly Glu Thr Val Lys His Glu Ile Ser Tyr Arg Ser Arg
340 345 350
Arg Leu Glu Gln Thr Ser Val Arg Asp Pro Ser Pro Glu Ala Asp Ala
355 360 365
Pro Val Leu Gly Ser Pro Glu Lys Glu Glu Ala Ala Ser Glu Pro Pro
370 375 380
Ala Ala Ala Pro Asp Ala Ala Pro Pro Pro Pro Asp Arg Pro Ile Glu
385 390 395 400
Lys Lys Ser Tyr Ser Arg Ala Arg Arg Thr Arg Thr Lys Val Gly Asp
405 410 415
Ala Val Lys Leu Ala Glu Glu Val Pro Pro Pro Pro Glu Gly Leu Ile
420 425 430
Pro Ala Pro Pro Val Pro Glu Thr Thr Pro Thr Pro Pro Thr Lys Thr
435 440 445
Gly Thr Trp Glu Ala Pro Val Asp Ser Ser Thr Ser Gly Leu Glu Gln
450 455 460
Asp Val Ala Gln Leu Asn Ile Ala Glu Gln Asn Trp Ser Pro Gly Gln
465 470 475 480
Pro Ser Phe Leu Gln Pro Arg Glu Leu Arg Gly Met Pro Asn His Ile
485 490 495
His Met Gly Ala Gly Pro Pro Pro Gln Phe Asn Arg Met Glu Glu Met
500 505 510
Gly Val Gln Gly Gly Arg Ala Lys Arg Tyr Ser Ser Gln Arg Gln Arg
515 520 525
Pro Val Pro Glu Pro Pro Ala Pro Pro Val His Ile Ser Ile Met Glu
530 535 540
Gly His Tyr Tyr Asp Pro Leu Gln Phe Gln Gly Pro Ile Tyr Thr His
545 550 555 560
Gly Asp Ser Pro Ala Pro Leu Pro Pro Gln Gly Met Leu Val Gln Pro
565 570 575
Gly Met Asn Leu Pro His Pro Gly Leu His Pro His Gln Thr Pro Ala
580 585 590
Pro Leu Pro Asn Pro Gly Leu Tyr Pro Pro Pro Val Ser Met Ser Pro
595 600 605
Gly Gln Pro Pro Pro Gln Gln Leu Leu Ala Pro Thr Tyr Phe Ser Ala
610 615 620
Pro Gly Val Met Asn Phe Gly Asn Pro Ser Tyr Pro Tyr Ala Pro Gly
625 630 635 640
Ala Leu Pro Pro Pro Pro Pro Pro His Leu Tyr Pro Asn Thr Gln Ala
645 650 655
Pro Ser Gln Val Tyr Gly Gly Val Thr Tyr Tyr Asn Pro Ala Gln Gln
660 665 670
Gln Val Gln Pro Lys Pro Ser Pro Pro Arg Arg Thr Pro Gln Pro Val
675 680 685
Thr Ile Lys Pro Pro Pro Pro Glu Val Val Ser Arg Gly Ser Ser
690 695 700

<210> 44

<211> 560

<212> PRT

<213> Homo sapiens

<400> 44
Met Pro Gln Thr Arg Ser Gln Ala Gln Ala Thr Ile Ser Phe Pro Lys
1 5 10 15
Arg Lys Leu Ser Arg Ala Leu Asn Lys Ala Lys Asn Ser Ser Asp Ala
20 25 30
Lys Leu Glu Pro Thr Asn Val Gln Thr Val Thr Cys Ser Pro Arg Val
35 40 45

Lys Ala Leu Pro Leu Ser Pro Arg Lys Arg Leu Gly Asp Asp Asn Leu
50 55 60
Cys Asn Thr Pro His Leu Pro Pro Cys Ser Pro Pro Lys Gln Gly Lys
65 70 75 80
Lys Glu Asn Gly Pro Pro His Ser His Thr Leu Lys Gly Arg Arg Leu
85 90 95
Val Phe Asp Asn Gln Leu Thr Ile Lys Ser Pro Ser Lys Arg Glu Leu
100 105 110
Ala Lys Val His Gln Asn Lys Ile Leu Ser Ser Val Arg Lys Ser Gln
115 120 125
Glu Ile Thr Thr Asn Ser Glu Gln Arg Cys Pro Leu Lys Lys Glu Ser
130 135 140
Ala Cys Val Arg Leu Phe Lys Gln Glu Gly Thr Cys Tyr Gln Gln Ala
145 150 155 160
Lys Leu Val Leu Asn Thr Ala Val Pro Asp Arg Leu Pro Ala Arg Glu
165 170 175
Arg Glu Met Asp Val Ile Arg Asn Phe Leu Arg Glu His Ile Cys Gly
180 185 190
Lys Lys Ala Gly Ser Leu Tyr Leu Ser Gly Ala Pro Gly Thr Gly Lys
195 200 205
Thr Ala Cys Leu Ser Arg Ile Leu Gln Asp Leu Lys Lys Glu Leu Lys
210 215 220
Gly Phe Lys Thr Ile Met Leu Asn Cys Met Ser Leu Arg Thr Ala Gln
225 230 235 240
Ala Val Phe Pro Ala Ile Ala Gln Glu Ile Cys Gln Glu Glu Val Ser
245 250 255
Arg Pro Ala Gly Lys Asp Met Met Arg Lys Leu Glu Lys His Met Thr
260 265 270
Ala Glu Lys Gly Pro Met Ile Val Leu Val Leu Asp Glu Met Asp Gln
275 280 285
Leu Asp Ser Lys Gly Gln Asp Val Leu Tyr Thr Leu Phe Glu Trp Pro
290 295 300
Trp Leu Ser Asn Ser His Leu Val Leu Ile Gly Ile Ala Asn Thr Leu
305 310 315 320
Asp Leu Thr Asp Arg Ile Leu Pro Arg Leu Gln Ala Arg Glu Lys Cys
325 330 335
Lys Pro Gln Leu Leu Asn Phe Pro Pro Tyr Thr Arg Asn Gln Ile Val
340 345 350
Thr Ile Leu Gln Asp Arg Leu Asn Gln Val Ser Arg Asp Gln Val Leu
355 360 365
Asp Asn Ala Ala Val Gln Phe Cys Ala Arg Lys Val Ser Ala Val Ser
370 375 380
Gly Asp Val Arg Lys Ala Leu Asp Val Cys Arg Arg Ala Ile Glu Ile
385 390 395 400
Val Glu Ser Asp Val Lys Ser Gln Thr Ile Leu Lys Pro Leu Ser Glu
405 410 415
Cys Lys Ser Pro Ser Glu Pro Leu Ile Pro Lys Arg Val Gly Leu Ile
420 425 430
His Ile Ser Gln Val Ile Ser Glu Val Asp Gly Asn Arg Met Thr Leu
435 440 445
Ser Gln Glu Gly Ala Gln Asp Ser Phe Pro Leu Gln Gln Lys Ile Leu
450 455 460
Val Cys Ser Leu Met Leu Leu Ile Arg Gln Leu Lys Ile Lys Glu Val
465 470 475 480
Thr Leu Gly Lys Leu Tyr Glu Ala Tyr Ser Lys Val Cys Arg Lys Gln
485 490 495
Gln Val Ala Ala Val Asp Gln Ser Glu Cys Leu Ser Leu Ser Gly Leu
500 505 510
Leu Glu Ala Arg Gly Ile Leu Gly Leu Lys Arg Asn Lys Glu Thr Arg
515 520 525
Leu Thr Lys Val Phe Phe Lys Ile Glu Glu Lys Glu Ile Glu His Ala
530 535 540
Leu Lys Asp Lys Ala Leu Ile Gly Asn Ile Leu Ala Thr Gly Leu Pro
545 550 555 560
<210> 45

<211> 462

<212> PRT

<213> Homo sapiens

<400> 45

```
Met Ala Ser Asn Ser Ser Ser Cys Pro Thr Pro Gly Gly Gly His Leu
1               5                   10                  15
Asn Gly Tyr Pro Val Pro Pro Tyr Ala Phe Phe Phe Pro Pro Met Leu
            20                  25                  30
Gly Gly Leu Ser Pro Pro Gly Ala Leu Thr Thr Leu Gln His Gln Leu
        35                  40                  45
Pro Val Ser Gly Tyr Ser Thr Pro Ser Pro Ala Thr Ile Glu Thr Gln
    50                  55                  60
Ser Ser Ser Ser Glu Glu Ile Val Pro Ser Pro Pro Ser Pro Pro Pro
65                  70                  75                  80
Leu Pro Arg Ile Tyr Lys Pro Cys Phe Val Cys Gln Asp Lys Ser Ser
                85                  90                  95
Gly Tyr His Tyr Gly Val Ser Ala Cys Glu Gly Cys Lys Gly Phe Phe
            100                 105                 110
Arg Arg Ser Ile Gln Lys Asn Met Val Tyr Thr Cys His Arg Asp Lys
        115                 120                 125
Asn Cys Ile Ile Asn Lys Val Thr Arg Asn Arg Cys Gln Tyr Cys Arg
    130                 135                 140
Leu Gln Lys Cys Phe Glu Val Gly Met Ser Lys Glu Ser Val Arg Asn
145                 150                 155                 160
Asp Arg Asn Lys Lys Lys Lys Glu Val Pro Lys Pro Glu Cys Ser Glu
                165                 170                 175
Ser Tyr Thr Leu Thr Pro Glu Val Gly Glu Leu Ile Glu Lys Val Arg
            180                 185                 190
Lys Ala His Gln Glu Thr Phe Pro Ala Leu Cys Gln Leu Gly Lys Tyr
        195                 200                 205
Thr Thr Asn Asn Ser Ser Glu Gln Arg Val Ser Leu Asp Ile Asp Leu
    210                 215                 220
Trp Asp Lys Phe Ser Glu Leu Ser Thr Lys Cys Ile Ile Lys Thr Val
225                 230                 235                 240
Glu Phe Ala Lys Gln Leu Pro Gly Phe Thr Thr Leu Thr Ile Ala Asp
                245                 250                 255
Gln Ile Thr Leu Leu Lys Ala Ala Cys Leu Asp Ile Leu Ile Leu Arg
            260                 265                 270
Ile Cys Thr Arg Tyr Thr Pro Glu Gln Asp Thr Met Thr Phe Ser Asp
        275                 280                 285
Gly Leu Thr Leu Asn Arg Thr Gln Met His Asn Ala Gly Phe Gly Pro
    290                 295                 300
Leu Thr Asp Leu Val Phe Ala Phe Ala Asn Gln Leu Leu Pro Leu Glu
305                 310                 315                 320
Met Asp Asp Ala Glu Thr Gly Leu Leu Ser Ala Ile Cys Leu Ile Cys
                325                 330                 335
Gly Asp Arg Gln Asp Leu Glu Gln Pro Asp Arg Val Asp Met Leu Gln
            340                 345                 350
Glu Pro Leu Leu Glu Ala Leu Lys Val Tyr Val Arg Lys Arg Arg Pro
        355                 360                 365
Ser Arg Pro His Met Phe Pro Lys Met Leu Met Lys Ile Thr Asp Leu
    370                 375                 380
Arg Ser Ile Ser Ala Lys Gly Ala Glu Arg Val Ile Thr Leu Lys Met
385                 390                 395                 400
Glu Ile Pro Gly Ser Met Pro Pro Leu Ile Gln Glu Met Leu Glu Asn
                405                 410                 415
Ser Glu Gly Leu Asp Thr Leu Ser Gly Gln Pro Gly Gly Gly Gly Arg
            420                 425                 430
Asp Gly Gly Gly Leu Ala Pro Pro Pro Gly Ser Cys Ser Pro Ser Leu
        435                 440                 445
```

Ser Pro Ser Ser Asn Arg Ser Ser Pro Ala Thr His Ser Pro
450 455 460

<210> 46

<211> 1531

<212> PRT

<213> Homo sapiens

<400> 46

Met Glu Val Ser Pro Leu Gln Pro Val Asn Glu Asn Met Gln Val Asn
1 5 10 15
Lys Ile Lys Lys Asn Glu Asp Ala Lys Lys Arg Leu Ser Val Glu Arg
20 25 30
Ile Tyr Gln Lys Lys Thr Gln Leu Glu His Ile Leu Leu Arg Pro Asp
35 40 45
Thr Tyr Ile Gly Ser Val Glu Leu Val Thr Gln Gln Met Trp Val Tyr
50 55 60
Asp Glu Asp Val Gly Ile Asn Tyr Arg Glu Val Thr Phe Val Pro Gly
65 70 75 80
Leu Tyr Lys Ile Phe Asp Glu Ile Leu Val Asn Ala Ala Asp Asn Lys
85 90 95
Gln Arg Asp Pro Lys Met Ser Cys Ile Arg Val Thr Ile Asp Pro Glu
100 105 110
Asn Asn Leu Ile Ser Ile Trp Asn Asn Gly Lys Gly Ile Pro Val Val
115 120 125
Glu His Lys Val Glu Lys Met Tyr Val Pro Ala Leu Ile Phe Gly Gln
130 135 140
Leu Leu Thr Ser Ser Asn Tyr Asp Asp Asp Glu Lys Lys Val Thr Gly
145 150 155 160
Gly Arg Asn Gly Tyr Gly Ala Lys Leu Cys Asn Ile Phe Ser Thr Lys
165 170 175
Phe Thr Val Glu Thr Ala Ser Arg Glu Tyr Lys Lys Met Phe Lys Gln
180 185 190
Thr Trp Met Asp Asn Met Gly Arg Ala Gly Glu Met Glu Leu Lys Pro
195 200 205
Phe Asn Gly Glu Asp Tyr Thr Cys Ile Thr Phe Gln Pro Asp Leu Ser
210 215 220
Lys Phe Lys Met Gln Ser Leu Asp Lys Asp Ile Val Ala Leu Met Val
225 230 235 240
Arg Arg Ala Tyr Asp Ile Ala Gly Ser Thr Lys Asp Val Lys Val Phe
245 250 255
Leu Asn Gly Asn Lys Leu Pro Val Lys Gly Phe Arg Ser Tyr Val Asp
260 265 270
Met Tyr Leu Lys Asp Lys Leu Asp Glu Thr Gly Asn Ser Leu Lys Val
275 280 285
Ile His Glu Gln Val Asn His Arg Trp Glu Val Cys Leu Thr Met Ser
290 295 300
Glu Lys Gly Phe Gln Gln Ile Ser Phe Val Asn Ser Ile Ala Thr Ser
305 310 315 320
Lys Gly Gly Arg His Val Asp Tyr Val Ala Asp Gln Ile Val Thr Lys
325 330 335
Leu Val Asp Val Val Lys Lys Lys Asn Lys Gly Gly Val Ala Val Lys
340 345 350
Ala His Gln Val Lys Asn His Met Trp Ile Phe Val Asn Ala Leu Ile
355 360 365
Glu Asn Pro Thr Phe Asp Ser Gln Thr Lys Glu Asn Met Thr Leu Gln
370 375 380
Pro Lys Ser Phe Gly Ser Thr Cys Gln Leu Ser Glu Lys Phe Ile Lys
385 390 395 400
Ala Ala Ile Gly Cys Gly Ile Val Glu Ser Ile Leu Asn Trp Val Lys
405 410 415

Phe Lys Ala Gln Val Gln Leu Asn Lys Lys Cys Ser Ala Val Lys His
420 425 430
Asn Arg Ile Lys Gly Ile Pro Lys Leu Asp Asp Ala Asn Asp Ala Gly
435 440 445
Gly Arg Asn Ser Thr Glu Cys Thr Leu Ile Leu Thr Glu Gly Asp Ser
450 455 460
Ala Lys Thr Leu Ala Val Ser Gly Leu Gly Val Val Gly Arg Asp Lys
465 470 475 480
Tyr Gly Val Phe Pro Leu Arg Gly Lys Ile Leu Asn Val Arg Glu Ala
485 490 495
Ser His Lys Gln Ile Met Glu Asn Ala Glu Ile Asn Asn Ile Ile Lys
500 505 510
Ile Val Gly Leu Gln Tyr Lys Lys Asn Tyr Glu Asp Glu Asp Ser Leu
515 520 525
Lys Thr Leu Arg Tyr Gly Lys Ile Met Ile Met Thr Asp Gln Asp Gln
530 535 540
Asp Gly Ser His Ile Lys Gly Leu Leu Ile Asn Phe Ile His His Asn
545 550 555 560
Trp Pro Ser Leu Leu Arg His Arg Phe Leu Glu Glu Phe Ile Thr Pro
565 570 575
Ile Val Lys Val Ser Lys Asn Lys Gln Glu Met Ala Phe Tyr Ser Leu
580 585 590
Pro Glu Phe Glu Glu Trp Lys Ser Ser Thr Pro Asn His Lys Lys Trp
595 600 605
Lys Val Lys Tyr Tyr Lys Gly Leu Gly Thr Ser Thr Ser Lys Glu Ala
610 615 620
Lys Glu Tyr Phe Ala Asp Met Lys Arg His Arg Ile Gln Phe Lys Tyr
625 630 635 640
Ser Gly Pro Glu Asp Asp Ala Ala Ile Ser Leu Ala Phe Ser Lys Lys
645 650 655
Gln Ile Asp Asp Arg Lys Glu Trp Leu Thr Asn Phe Met Glu Asp Arg
660 665 670
Arg Gln Arg Lys Leu Leu Gly Leu Pro Glu Asp Tyr Leu Tyr Gly Gln
675 680 685
Thr Thr Thr Tyr Leu Thr Tyr Asn Asp Phe Ile Asn Lys Glu Leu Ile
690 695 700
Leu Phe Ser Asn Ser Asp Asn Glu Arg Ser Ile Pro Ser Met Val Asp
705 710 715 720
Gly Leu Lys Pro Gly Gln Arg Lys Val Leu Phe Thr Cys Phe Lys Arg
725 730 735
Asn Asp Lys Arg Glu Val Lys Val Ala Gln Leu Ala Gly Ser Val Ala
740 745 750
Glu Met Ser Ser Tyr His His Gly Glu Met Ser Leu Met Met Thr Ile
755 760 765
Ile Asn Leu Ala Gln Asn Phe Val Gly Ser Asn Asn Leu Asn Leu Leu
770 775 780
Gln Pro Ile Gly Gln Phe Gly Thr Arg Leu His Gly Gly Lys Asp Ser
785 790 795 800
Ala Ser Pro Arg Tyr Ile Phe Thr Met Leu Ser Ser Leu Ala Arg Leu
805 810 815
Leu Phe Pro Pro Lys Asp Asp His Thr Leu Lys Phe Leu Tyr Asp Asp
820 825 830
Asn Gln Arg Val Glu Pro Glu Trp Tyr Ile Pro Ile Ile Pro Met Val
835 840 845
Leu Ile Asn Gly Ala Glu Gly Ile Gly Thr Gly Trp Ser Cys Lys Ile
850 855 860
Pro Asn Phe Asp Val Arg Glu Ile Val Asn Asn Ile Arg Arg Leu Met
865 870 875 880
Asp Gly Glu Glu Pro Leu Pro Met Leu Pro Ser Tyr Lys Asn Phe Lys
885 890 895
Gly Thr Ile Glu Glu Leu Ala Pro Asn Gln Tyr Val Ile Ser Gly Glu
900 905 910
Val Ala Ile Leu Asn Ser Thr Thr Ile Glu Ile Ser Glu Leu Pro Val
915 920 925
Arg Thr Trp Thr Gln Thr Tyr Lys Glu Gln Val Leu Glu Pro Met Leu
930 935 940

Asn Gly Thr Glu Lys Thr Pro Pro Leu Ile Thr Asp Tyr Arg Glu Tyr
945 950 955 960
His Thr Asp Thr Thr Val Lys Phe Val Val Lys Met Thr Glu Glu Lys
965 970 975
Leu Ala Glu Ala Glu Arg Val Gly Leu His Lys Val Phe Lys Leu Gln
980 985 990
Thr Ser Leu Thr Cys Asn Ser Met Val Leu Phe Asp His Val Gly Cys
995 1000 1005
Leu Lys Lys Tyr Asp Thr Val Leu Asp Ile Leu Arg Asp Phe Phe
1010 1015 1020
Glu Leu Arg Leu Lys Tyr Tyr Gly Leu Arg Lys Glu Trp Leu Leu
1025 1030 1035
Gly Met Leu Gly Ala Glu Ser Ala Lys Leu Asn Asn Gln Ala Arg
1040 1045 1050
Phe Ile Leu Glu Lys Ile Asp Gly Lys Ile Ile Ile Glu Asn Lys
1055 1060 1065
Pro Lys Lys Glu Leu Ile Lys Val Leu Ile Gln Arg Gly Tyr Asp
1070 1075 1080
Ser Asp Pro Val Lys Ala Trp Lys Glu Ala Gln Gln Lys Val Pro
1085 1090 1095
Asp Glu Glu Glu Asn Glu Glu Ser Asp Asn Glu Lys Glu Thr Glu
1100 1105 1110
Lys Ser Asp Ser Val Thr Asp Ser Gly Pro Thr Phe Asn Tyr Leu
1115 1120 1125
Leu Asp Met Pro Leu Trp Tyr Leu Thr Lys Glu Lys Lys Asp Glu
1130 1135 1140
Leu Cys Arg Leu Arg Asn Glu Lys Glu Gln Glu Leu Asp Thr Leu
1145 1150 1155
Lys Arg Lys Ser Pro Ser Asp Leu Trp Lys Glu Asp Leu Ala Thr
1160 1165 1170
Phe Ile Glu Glu Leu Glu Ala Val Glu Ala Lys Glu Lys Gln Asp
1175 1180 1185
Glu Gln Val Gly Leu Pro Gly Lys Gly Gly Lys Ala Lys Gly Lys
1190 1195 1200
Lys Thr Gln Met Ala Glu Val Leu Pro Ser Pro Arg Gly Gln Arg
1205 1210 1215
Val Ile Pro Arg Ile Thr Ile Glu Met Lys Ala Glu Ala Glu Lys
1220 1225 1230
Lys Asn Lys Lys Lys Ile Lys Asn Glu Asn Thr Glu Gly Ser Pro
1235 1240 1245
Gln Glu Asp Gly Val Glu Leu Glu Gly Leu Lys Gln Arg Leu Glu
1250 1255 1260
Lys Lys Gln Lys Arg Glu Pro Gly Thr Lys Thr Lys Lys Gln Thr
1265 1270 1275
Thr Leu Ala Phe Lys Pro Ile Lys Lys Gly Lys Lys Arg Asn Pro
1280 1285 1290
Trp Ser Asp Ser Glu Ser Asp Arg Ser Ser Asp Glu Ser Asn Phe
1295 1300 1305
Asp Val Pro Pro Arg Glu Thr Glu Pro Arg Arg Ala Ala Thr Lys
1310 1315 1320
Thr Lys Phe Thr Met Asp Leu Asp Ser Asp Glu Asp Phe Ser Asp
1325 1330 1335
Phe Asp Glu Lys Thr Asp Asp Glu Asp Phe Val Pro Ser Asp Ala
1340 1345 1350
Ser Pro Pro Lys Thr Lys Thr Ser Pro Lys Leu Ser Asn Lys Glu
1355 1360 1365
Leu Lys Pro Gln Lys Ser Val Val Ser Asp Leu Glu Ala Asp Asp
1370 1375 1380
Val Lys Gly Ser Val Pro Leu Ser Ser Ser Pro Pro Ala Thr His
1385 1390 1395
Phe Pro Asp Glu Thr Glu Ile Thr Asn Pro Val Pro Lys Lys Asn
1400 1405 1410
Val Thr Val Lys Lys Thr Ala Ala Lys Ser Gln Ser Ser Thr Ser
1415 1420 1425
Thr Thr Gly Ala Lys Lys Arg Ala Ala Pro Lys Gly Thr Lys Arg
1430 1435 1440

Asp Pro Ala Leu Asn Ser Gly Val Ser Gln Lys Pro Asp Pro Ala
1445 1450 1455
Lys Thr Lys Asn Arg Arg Lys Arg Lys Pro Ser Thr Ser Asp Asp
1460 1465 1470
Ser Asp Ser Asn Phe Glu Lys Ile Val Ser Lys Ala Val Thr Ser
1475 1480 1485
Lys Lys Ser Lys Gly Glu Ser Asp Asp Phe His Met Asp Phe Asp
1490 1495 1500
Ser Ala Val Ala Pro Arg Ala Lys Ser Val Arg Ala Lys Lys Pro
1505 1510 1515
Ile Lys Tyr Leu Glu Glu Ser Asp Glu Asp Asp Leu Phe
1520 1525 1530

<210> 47

<211> 258

<212> PRT

<213> Homo sapiens

<400> 47
Met Leu Pro Leu Cys Leu Val Ala Ala Leu Leu Leu Ala Ala Gly Pro
1 5 10 15
Gly Pro Ser Leu Gly Asp Glu Ala Ile His Cys Pro Pro Cys Ser Glu
20 25 30
Glu Lys Leu Ala Arg Cys Arg Pro Pro Val Gly Cys Glu Glu Leu Val
35 40 45
Arg Glu Pro Gly Cys Gly Cys Cys Ala Thr Cys Ala Leu Gly Leu Gly
50 55 60
Met Pro Cys Gly Val Tyr Thr Pro Arg Cys Gly Ser Gly Leu Arg Cys
65 70 75 80
Tyr Pro Pro Arg Gly Val Glu Lys Pro Leu His Thr Leu Met His Gly
85 90 95
Gln Gly Val Cys Met Glu Leu Ala Glu Ile Glu Ala Ile Gln Glu Ser
100 105 110
Leu Gln Pro Ser Asp Lys Asp Glu Gly Asp His Pro Asn Asn Ser Phe
115 120 125
Ser Pro Cys Ser Ala His Asp Arg Arg Cys Leu Gln Lys His Phe Ala
130 135 140
Lys Ile Arg Asp Arg Ser Thr Ser Gly Gly Lys Met Lys Val Asn Gly
145 150 155 160
Ala Pro Arg Glu Asp Ala Arg Pro Val Pro Gln Gly Ser Cys Gln Ser
165 170 175
Glu Leu His Arg Ala Leu Glu Arg Leu Ala Ala Ser Gln Ser Arg Thr
180 185 190
His Glu Asp Leu Tyr Ile Ile Pro Ile Pro Asn Cys Asp Arg Asn Gly
195 200 205
Asn Phe His Pro Lys Gln Cys His Pro Ala Leu Asp Gly Gln Arg Gly
210 215 220
Lys Cys Trp Cys Val Asp Arg Lys Thr Gly Val Lys Leu Pro Gly Gly
225 230 235 240
Leu Glu Pro Lys Gly Glu Leu Asp Cys His Gln Leu Ala Asp Ser Phe
245 250 255
Arg Glu

<210> 48

<211> 378

<212> PRT

<213> Homo sapiens

<400> 48
Met Asp Leu Gly Lys Pro Met Lys Ser Val Leu Val Val Ala Leu Leu
1 5 10 15
Val Ile Phe Gln Val Cys Leu Cys Gln Asp Glu Val Thr Asp Asp Tyr
20 25 30
Ile Gly Asp Asn Thr Thr Val Asp Tyr Thr Leu Phe Glu Ser Leu Cys
35 40 45
Ser Lys Lys Asp Val Arg Asn Phe Lys Ala Trp Phe Leu Pro Ile Met
50 55 60
Tyr Ser Ile Ile Cys Phe Val Gly Leu Leu Gly Asn Gly Leu Val Val
65 70 75 80
Leu Thr Tyr Ile Tyr Phe Lys Arg Leu Lys Thr Met Thr Asp Thr Tyr
85 90 95
Leu Leu Asn Leu Ala Val Ala Asp Ile Leu Phe Leu Leu Thr Leu Pro
100 105 110
Phe Trp Ala Tyr Ser Ala Ala Lys Ser Trp Val Phe Gly Val His Phe
115 120 125
Cys Lys Leu Ile Phe Ala Ile Tyr Lys Met Ser Phe Phe Ser Gly Met
130 135 140
Leu Leu Leu Leu Cys Ile Ser Ile Asp Arg Tyr Val Ala Ile Val Gln
145 150 155 160
Ala Val Ser Ala His Arg His Arg Ala Arg Val Leu Leu Ile Ser Lys
165 170 175
Leu Ser Cys Val Gly Ile Trp Ile Leu Ala Thr Val Leu Ser Ile Pro
180 185 190
Glu Leu Leu Tyr Ser Asp Leu Gln Arg Ser Ser Ser Glu Gln Ala Met
195 200 205
Arg Cys Ser Leu Ile Thr Glu His Val Glu Ala Phe Ile Thr Ile Gln
210 215 220
Val Ala Gln Met Val Ile Gly Phe Leu Val Pro Leu Leu Ala Met Ser
225 230 235 240
Phe Cys Tyr Leu Val Ile Ile Arg Thr Leu Leu Gln Ala Arg Asn Phe
245 250 255
Glu Arg Asn Lys Ala Ile Lys Val Ile Ile Ala Val Val Val Val Phe
260 265 270
Ile Val Phe Gln Leu Pro Tyr Asn Gly Val Val Leu Ala Gln Thr Val
275 280 285
Ala Asn Phe Asn Ile Thr Ser Ser Thr Cys Glu Leu Ser Lys Gln Leu
290 295 300
Asn Ile Ala Tyr Asp Val Thr Tyr Ser Leu Ala Cys Val Arg Cys Cys
305 310 315 320
Val Asn Pro Phe Leu Tyr Ala Phe Ile Gly Val Lys Phe Arg Asn Asp
325 330 335
Leu Phe Lys Leu Phe Lys Asp Leu Gly Cys Leu Ser Gln Glu Gln Leu
340 345 350
Arg Gln Trp Ser Ser Cys Arg His Ile Arg Arg Ser Ser Met Ser Val
355 360 365
Glu Ala Glu Thr Thr Thr Thr Phe Ser Pro
370 375

<210> 49

<211> 411

<212> PRT

<213> Homo sapiens

<400> 49
Met Ser Lys Arg Pro Ser Tyr Ala Pro Pro Pro Thr Pro Ala Pro Ala
1 5 10 15
Thr Gln Met Pro Ser Thr Pro Gly Phe Val Gly Tyr Asn Pro Tyr Ser
20 25 30

His Leu Ala Tyr Asn Asn Tyr Arg Leu Gly Gly Asn Pro Ser Thr Asn
35 40 45
Ser Arg Val Thr Ala Ser Ser Gly Ile Thr Ile Pro Lys Pro Pro Lys
50 55 60
Pro Pro Asp Lys Pro Leu Met Pro Tyr Met Arg Tyr Ser Arg Lys Val
65 70 75 80
Trp Asp Gln Val Lys Ala Ser Asn Pro Asp Leu Lys Leu Trp Glu Ile
85 90 95
Gly Lys Ile Ile Gly Gly Met Trp Arg Asp Leu Thr Asp Glu Glu Lys
100 105 110
Gln Glu Tyr Leu Asn Glu Tyr Glu Ala Glu Lys Ile Glu Tyr Asn Glu
115 120 125
Ser Met Lys Ala Tyr His Asn Ser Pro Ala Tyr Leu Ala Tyr Ile Asn
130 135 140
Ala Lys Ser Arg Ala Glu Ala Ala Leu Glu Glu Glu Ser Arg Gln Arg
145 150 155 160
Gln Ser Arg Met Glu Lys Gly Glu Pro Tyr Met Ser Ile Gln Pro Ala
165 170 175
Glu Asp Pro Asp Asp Tyr Asp Asp Gly Phe Ser Met Lys His Thr Ala
180 185 190
Thr Ala Arg Phe Gln Arg Asn His Arg Leu Ile Ser Glu Ile Leu Ser
195 200 205
Glu Ser Val Val Pro Asp Val Arg Ser Val Val Thr Thr Ala Arg Met
210 215 220
Gln Val Leu Lys Arg Gln Val Gln Ser Leu Met Val His Gln Arg Lys
225 230 235 240
Leu Glu Ala Glu Leu Leu Gln Ile Glu Glu Arg His Gln Glu Lys Lys
245 250 255
Arg Lys Phe Leu Glu Ser Thr Asp Ser Phe Asn Asn Glu Leu Lys Arg
260 265 270
Leu Cys Gly Leu Lys Val Glu Val Asp Met Glu Lys Ile Ala Ala Glu
275 280 285
Ile Ala Gln Ala Glu Glu Gln Ala Arg Lys Arg Gln Glu Glu Arg Glu
290 295 300
Lys Glu Ala Ala Glu Gln Ala Glu Arg Ser Gln Ser Ser Ile Val Pro
305 310 315 320
Glu Glu Glu Gln Ala Ala Asn Lys Gly Glu Glu Lys Lys Asp Asp Glu
325 330 335
Asn Ile Pro Met Glu Thr Glu Glu Thr His Leu Glu Glu Thr Thr Glu
340 345 350
Ser Gln Gln Asn Gly Glu Glu Gly Thr Ser Thr Pro Glu Asp Lys Glu
355 360 365
Ser Gly Gln Glu Gly Val Asp Ser Met Ala Glu Glu Gly Thr Ser Asp
370 375 380
Ser Asn Thr Gly Ser Glu Ser Asn Ser Ala Thr Val Glu Glu Pro Pro
385 390 395 400
Thr Asp Pro Ile Pro Glu Asp Glu Lys Lys Glu
405 410

<210> 50

<211> 593

<212> PRT

<213> Homo sapiens

<400> 50
Met Ser Val Arg Tyr Ser Ser Ser Lys His Tyr Ser Ser Ser Arg Ser
1 5 10 15
Gly Gly Gly Gly Gly Gly Gly Gly Cys Gly Gly Gly Gly Gly Val Ser
20 25 30
Ser Leu Arg Ile Ser Ser Ser Lys Gly Ser Leu Gly Gly Gly Phe Ser
35 40 45

Ser Gly Gly Phe Ser Gly Gly Ser Phe Ser Arg Gly Ser Ser Gly Gly
50 55 60
Gly Cys Phe Gly Gly Ser Ser Gly Gly Tyr Gly Gly Leu Gly Gly Phe
65 70 75 80
Gly Gly Gly Ser Phe His Gly Ser Tyr Gly Ser Ser Ser Phe Gly Gly
85 90 95
Ser Tyr Gly Gly Ser Phe Gly Gly Gly Asn Phe Gly Gly Gly Ser Phe
100 105 110
Gly Gly Gly Ser Phe Gly Gly Gly Gly Phe Gly Gly Gly Gly Phe Gly
115 120 125
Gly Gly Phe Gly Gly Gly Phe Gly Gly Asp Gly Gly Leu Leu Ser Gly
130 135 140
Asn Glu Lys Val Thr Met Gln Asn Leu Asn Asp Arg Leu Ala Ser Tyr
145 150 155 160
Leu Asp Lys Val Arg Ala Leu Glu Glu Ser Asn Tyr Glu Leu Glu Gly
165 170 175
Lys Ile Lys Glu Trp Tyr Glu Lys His Gly Asn Ser His Gln Gly Glu
180 185 190
Pro Arg Asp Tyr Ser Lys Tyr Tyr Lys Thr Ile Asp Asp Leu Lys Asn
195 200 205
Gln Ile Leu Asn Leu Thr Thr Asp Asn Ala Asn Ile Leu Leu Gln Ile
210 215 220
Asp Asn Ala Arg Leu Ala Ala Asp Asp Phe Arg Leu Lys Tyr Glu Asn
225 230 235 240
Glu Val Ala Leu Arg Gln Ser Val Glu Ala Asp Ile Asn Gly Leu Arg
245 250 255
Arg Val Leu Asp Glu Leu Thr Leu Thr Lys Ala Asp Leu Glu Met Gln
260 265 270
Ile Glu Ser Leu Thr Glu Glu Leu Ala Tyr Leu Lys Lys Asn His Glu
275 280 285
Glu Glu Met Lys Asp Leu Arg Asn Val Ser Thr Gly Asp Val Asn Val
290 295 300
Glu Met Asn Ala Ala Pro Gly Val Asp Leu Thr Gln Leu Leu Asn Asn
305 310 315 320
Met Arg Ser Gln Tyr Glu Gln Leu Ala Glu Gln Asn Arg Lys Asp Ala
325 330 335
Glu Ala Trp Phe Asn Glu Lys Ser Lys Glu Leu Thr Thr Glu Ile Asp
340 345 350
Asn Asn Ile Glu Gln Ile Ser Ser Tyr Lys Ser Glu Ile Thr Glu Leu
355 360 365
Arg Arg Asn Val Gln Ala Leu Glu Ile Glu Leu Gln Ser Gln Leu Ala
370 375 380
Leu Lys Gln Ser Leu Glu Ala Ser Leu Ala Glu Thr Glu Gly Arg Tyr
385 390 395 400
Cys Val Gln Leu Ser Gln Ile His Ala Gln Ile Ser Ala Leu Glu Glu
405 410 415
Gln Leu Gln Gln Ile Arg Ala Glu Thr Glu Cys Gln Asn Thr Glu Tyr
420 425 430
Gln Gln Leu Leu Asp Ile Lys Ile Arg Leu Glu Asn Glu Ile Gln Thr
435 440 445
Tyr Arg Ser Leu Leu Glu Gly Glu Gly Ser Ser Gly Gly Gly Gly Arg
450 455 460
Gly Gly Gly Ser Phe Gly Gly Gly Tyr Gly Gly Gly Ser Ser Gly Gly
465 470 475 480
Gly Ser Ser Gly Gly Gly Tyr Gly Gly Gly His Gly Gly Ser Ser Gly
485 490 495
Gly Gly Tyr Gly Gly Gly Ser Ser Gly Gly Gly Ser Ser Gly Gly Gly
500 505 510
Tyr Gly Gly Gly Ser Ser Ser Gly Gly His Gly Gly Gly Ser Ser Ser
515 520 525
Gly Gly His Gly Gly Ser Ser Ser Gly Gly Tyr Gly Gly Gly Ser Ser
530 535 540
Gly Gly Gly Gly Gly Gly Tyr Gly Gly Gly Ser Ser Gly Gly Gly Ser
545 550 555 560
Ser Ser Gly Gly Gly Tyr Gly Gly Gly Ser Ser Ser Gly Gly His Lys
565 570 575

Ser Ser Ser Ser Gly Ser Val Gly Glu Ser Ser Ser Lys Gly Pro Arg
580 585 590

Tyr

<210> 51

<211> 494

<212> PRT

<213> Homo sapiens

<400> 51
Met Asp Leu Ser Asn Asn Thr Met Ser Leu Ser Val Arg Thr Pro Gly
1 5 10 15
Leu Ser Arg Arg Leu Ser Ser Gln Ser Val Ile Gly Arg Pro Arg Gly
20 25 30
Met Ser Ala Ser Ser Val Gly Ser Gly Tyr Gly Gly Ser Ala Phe Gly
35 40 45
Phe Gly Ala Ser Cys Gly Gly Gly Phe Ser Ala Ala Ser Met Phe Gly
50 55 60
Ser Ser Ser Gly Phe Gly Gly Gly Ser Gly Ser Ser Met Ala Gly Gly
65 70 75 80
Leu Gly Ala Gly Tyr Gly Arg Ala Leu Gly Gly Gly Ser Phe Gly Gly
85 90 95
Leu Gly Met Gly Phe Gly Gly Ser Pro Gly Gly Gly Ser Leu Gly Ile
100 105 110
Leu Ser Gly Asn Asp Gly Gly Leu Leu Ser Gly Ser Glu Lys Glu Thr
115 120 125
Met Gln Asn Leu Asn Asp Arg Leu Ala Ser Tyr Leu Asp Lys Val Arg
130 135 140
Ala Leu Glu Glu Ala Asn Thr Glu Leu Glu Asn Lys Ile Arg Glu Trp
145 150 155 160
Tyr Glu Thr Arg Gly Thr Gly Thr Ala Asp Ala Ser Gln Ser Asp Tyr
165 170 175
Ser Lys Tyr Tyr Pro Leu Ile Glu Asp Leu Arg Asn Lys Ile Ile Ser
180 185 190
Ala Ser Ile Gly Asn Ala Gln Leu Leu Leu Gln Ile Asp Asn Ala Arg
195 200 205
Leu Ala Ala Glu Asp Phe Arg Met Lys Tyr Glu Asn Glu Leu Ala Leu
210 215 220
Arg Gln Gly Val Glu Ala Asp Ile Asn Gly Leu Arg Arg Val Leu Asp
225 230 235 240
Glu Leu Thr Leu Thr Arg Thr Asp Leu Glu Met Gln Ile Glu Ser Leu
245 250 255
Asn Glu Glu Leu Ala Tyr Met Lys Lys Asn His Glu Asp Glu Leu Gln
260 265 270
Ser Phe Arg Val Gly Gly Pro Gly Glu Val Ser Val Glu Met Asp Ala
275 280 285
Ala Pro Gly Val Asp Leu Thr Arg Leu Leu Asn Asp Met Arg Ala Gln
290 295 300
Tyr Glu Thr Ile Ala Glu Gln Asn Arg Lys Asp Ala Glu Ala Trp Phe
305 310 315 320
Ile Glu Lys Ser Gly Glu Leu Arg Lys Glu Ile Ser Thr Asn Thr Glu
325 330 335
Gln Leu Gln Ser Ser Lys Ser Glu Val Thr Asp Leu Arg Arg Ala Phe
340 345 350
Gln Asn Leu Glu Ile Glu Leu Gln Ser Gln Leu Ala Met Lys Ser
355 360 365
Leu Glu Asp Ser Leu Ala Glu Ala Glu Gly Asp Tyr Cys Ala Gln Leu
370 375 380
Ser Gln Val Gln Gln Leu Ile Ser Asn Leu Glu Ala Gln Leu Leu Gln
385 390 395 400

Val Arg Ala Asp Ala Glu Arg Gln Asn Val Asp His Gln Arg Leu Leu
405 410 415
Asn Val Lys Ala Arg Leu Glu Leu Glu Ile Glu Thr Tyr Arg Arg Leu
420 425 430
Leu Asp Gly Glu Ala Gln Gly Asp Gly Leu Glu Glu Ser Leu Phe Val
435 440 445
Thr Asp Ser Lys Ser Gln Ala Gln Ser Thr Asp Ser Ser Lys Asp Pro
450 455 460
Thr Lys Thr Arg Lys Ile Lys Thr Val Val Gln Glu Met Val Asn Gly
465 470 475 480
Glu Val Val Ser Ser Gln Val Gln Glu Ile Glu Glu Leu Met
485 490

<210> 52

<211> 361

<212> PRT

<213> Homo sapiens

<400> 52
Cys Asn Trp Phe Cys Glu Gly Ser Phe Asn Gly Ser Glu Lys Glu Thr
1 5 10 15
Met Gln Phe Leu Asn Asp Arg Leu Ala Ser Tyr Leu Glu Lys Val Arg
20 25 30
His Val Glu Arg Asp Asn Ala Glu Leu Glu Asn Leu Ile Arg Glu Arg
35 40 45
Ser Gln Gln Gln Glu Pro Leu Leu Cys Pro Ser Tyr Gln Ser Tyr Phe
50 55 60
Lys Thr Ile Glu Glu Leu Gln Gln Lys Ile Leu Cys Ser Lys Ser Glu
65 70 75 80
Asn Ala Arg Leu Val Val Gln Ile Asp Asn Ala Lys Leu Ala Ala Asp
85 90 95
Asp Phe Arg Thr Lys Tyr Gln Thr Glu Gln Ser Leu Arg Gln Leu Val
100 105 110
Glu Ser Asp Ile Asn Ser Leu Arg Arg Ile Leu Asp Glu Leu Thr Leu
115 120 125
Cys Arg Ser Asp Leu Glu Ala Gln Met Glu Ser Leu Lys Glu Glu Leu
130 135 140
Leu Ser Leu Lys Gln Asn His Glu Gln Glu Val Asn Thr Leu Arg Cys
145 150 155 160
Gln Leu Gly Asp Arg Leu Asn Val Glu Val Asp Ala Ala Pro Ala Val
165 170 175
Asp Leu Asn Gln Val Leu Asn Glu Thr Arg Asn Gln Tyr Glu Ala Leu
180 185 190
Val Glu Thr Asn Arg Arg Glu Val Glu Gln Trp Phe Ala Thr Gln Thr
195 200 205
Glu Glu Leu Asn Lys Gln Val Val Ser Ser Ser Glu Gln Leu Gln Ser
210 215 220
Tyr Gln Ala Glu Ile Ile Glu Leu Arg Arg Thr Val Asn Ala Leu Glu
225 230 235 240
Ile Glu Leu Gln Ala Gln His Asn Leu Arg Tyr Ser Leu Glu Asn Thr
245 250 255
Leu Thr Glu Ser Glu Ala Arg Tyr Ser Ser Gln Leu Ser Gln Val Gln
260 265 270
Ser Leu Ile Thr Asn Val Glu Ser Gln Leu Ala Glu Ile Arg Ser Asp
275 280 285
Leu Glu Arg Gln Asn Gln Glu Tyr Gln Val Leu Leu Asp Val Arg Ala
290 295 300
Arg Leu Glu Cys Glu Ile Asn Thr Tyr Arg Ser Leu Leu Glu Ser Glu
305 310 315 320
Asp Cys Lys Leu Pro Ser Asn Pro Cys Ala Thr Thr Asn Ala Cys Glu
325 330 335

```
Lys Pro Ile Gly Ser Cys Val Thr Asn Pro Cys Gly Pro Arg Ser Arg
            340                 345                 350
Cys Gly Pro Cys Asn Thr Phe Gly Tyr
        355                 360
```

<210> 53

<211> 3282

<212> DNA

<213> Homo sapiens

<400> 53

```
atgaaggaga tggtaggagg ctgctgcgta tgttcggacg agaggggctg ggccgagaac    60
ccgctggtct actgcgatgg gcacgcgtgc agcgtggccg tccaccaagc ttgctatggc   120
atcgttcagg tgccaacggg accctggttc tgccggaaat gtgaatctca ggagcgagca   180
gccagggtga ggtgtgagct gtgcccacac aaagacgggg cattgaagag gactgataat   240
ggaggctggg cacacgtggt gtgtgccctc tacatccccg aggtgcaatt tgccaacgtg   300
ctcaccatgg agcccatcgt gctgcagtac gtgcctcatg atcgcttcaa caagacctgt   360
tacatctgcg aggagacggg ccgggagagc aaggcggcct cgggagcctg catgacctgt   420
aaccgccatg gatgtcgaca agctttccac gtcacctgtg cccaaatggc aggcttgctg   480
tgtgaggaag aagtgctgga ggtggacaac gtcaagtact gcggctactg caaataccac   540
ttcagcaaga tgaagacatc ccggcacagc agcgggggag gcggaggagg cgctggagga   600
ggaggtggca gcatgggggg aggtggcagt ggtttcatct ctgggaggag aagccggtca   660
gcctcaccat ccacgcagca ggagaagcac cccacccacc acgagagggg ccagaagaag   720
agtcgaaagg acaaagaacg ccttaagcag aagcacaaga agcggcctga gtcgcccccc   780
agcatcctca ccccgcccgt ggtccccact gctgacaagg tctcctcctc ggcttcctct   840
tcctcccacc acgaggccag cacgcaggag acctctgaga gcagcaggga gtcaaagggg   900
aaaaagtctt ccagccatag cctgagtcat aaagggaaga aactgagcag tgggaaaggt   960
gtgagcagtt ttacctccgc ctcctcttct tcctcctcct cttcctcctc ctctgggggg  1020
cccttccagc ctgcagtctc gtccctgcag agctcccctg acttctctgc attccccaag  1080
ctggagcagc cagaggagga caagtactcc aagcccacag cccccgcccc ttcagcccct  1140
ccttctccct cagctcccga gccccccaag gctgaccttt ttgagcagaa ggtggtcttc  1200
tctggctttg ggcccatcat gcgcttctcc accaccacct ccagctcagg ccgggcccgg  1260
gcgccctccc ctggggacta taagtctccc cacgtcacgg ggtctggggc ctcggcaggc  1320
acccacaaac ggatgcccgc actgagtgcc acccctgtgc ctgctgatga gacccctgag  1380
acaggcctga aggagaagaa gcacaaagcc agcaagagga gccgccatgg gccaggccgt  1440
cccaagggca gccggaacaa ggagggcact ggggggccag ctgccccatc cttgcccagt  1500
gcccagctgg ctggctttac cgccactgct gcctcaccct tctctggagg ttccctggtc  1560
agctccggcc tgggaggtct gtcctcccga acctttgggc cttctgggag cttgcccagc  1620
ttgagcctgg agtccccctt actaggggca ggcatctaca ccagtaataa ggaccccatc  1680
tcccacagtg gcgggatgct gcgggctgtc tgcagcaccc ctctctcctc cagcctcctg  1740
gggcccccag ggacctcggc cctgccccgc ctcagccgct ccccgttcac cagcaccctc  1800
ccctcctctt ctgcttctat ctccaccact caggtgtttt ctctggctgg ctctaccttt  1860
agcctcccct ctacccacat ctttggaacc cccatgggtg ccgttaatcc cctcctctcc  1920
caagctgaga gcagccacac agagccagac ctggaggact gcagcttccg gtgtcggggg  1980
acctcccctc aggagagtct gtcttccatg tcccccatca gcagcctccc cgcactcttc  2040
gaccagacag cctctgcacc ctgtgggggc ggccagttag acccggcggc cccagggacg  2100
actaacatgg agcagcttct ggagaagcag ggcgacgggg aggccggcgt caacatcgtg  2160
gagatgctga aggcgctgca cgcgctgcag aaggagaacc agcggctgca agagcagatc  2220
ctgagcctga cggccaaaaa ggagcggctg cagattctca acgtgcagct ctctgtgccc  2280
ttccctgccc tgcctgctgc cctgcctgcc gccaacggcc ctgtccctgg gccctatggc  2340
ctgcctcccc aagccgggag cagcgactcc ttgagcacca gcaagagccc tccgggaaag  2400
agcagcctcg gcctggacaa ctcgctgtcc acttcttctg aggacccaca ctcaggctgc  2460
ccgagccgca gcagctcgtc gctgtccttc cacagcacgc ccccaccgct gcccctcctc  2520
cagcagagcc ctgccactct gcccctggcc ctgcctgggg cccctgcccc actcccgccc  2580
cagccgcaga acgggttggg ccgggcaccc ggggcagcgg ggctgggggc catgcccatg  2640
gctgaggggc tgttgggggg gctggcaggc agtgggggcc tgcccctcaa tgggctcctt  2700
gggggggttga atggggccgc tgccccccaac cccgcaagct tgagccaggc tggcggggcc  2760
cccacgctgc agctgccagg ctgtctcaac agccttacag agcagcagag acatctcctt  2820
cagcagcaag agcagcagct ccagcaactc cagcagctcc tggcctcccc gcagctgacc  2880
ccggaacacc agactgttgt ctaccagatg atccagcaga tccagcagaa acgggagctg  2940
cagcgtctgc agatggctgg gggctcccag ctgcccatgg ccagcctgct ggcaggaagc  3000
tccaccccgc tgctgtctgc gggtacccct ggcctgctgc ccacagcgtc tgctccaccc  3060
```

```
ctgctgcccg ctggagccct agtggctccc tcgcttggca acaacacaag tctcatggcc   3120
gcagcagctg cagctgcagc agtagcagca gcaggcggac ctccagtcct cactgcccag   3180
accaacccct tcctcagcct gtcgggagca gagggcagtg gcggtggccc caaaggaggg   3240
accgctgaca aaggagcctc agccaaccag gaaaaaggct aa                      3282
```

<210> 54

<211> 2227

<212> DNA

<213> Homo sapiens

<400> 54

```
gagagcccga acaggaagag ggtacagctt tgtgcaggtc acatgcccac tgcagccctc     60
cagcctctgg tccccagagc ggactttgga agctgaactg cttttgttgc tggaagactt    120
atgttataat ttaccctggg tggaccaggg tcgtacaaaa gggcaacgct ccccagtccc    180
cccactcccg accccggaat catgcatcgg actacacgga tcaaaatcac agagctgaac    240
ccccacctca tgtgtgccct ctgcgggggg tacttcatcg acgccaccac tatcgtggag    300
tgcctgcatt ccttctgcaa aacctgcatc gtgcgctacc tggagaccaa caaatactgc    360
cccatgtgtg acgtgcaggt ccataaaacc cggccgctgc tgagcatcag gtctgacaaa    420
acacttcaag acattgtcta caaattggtc cctgggcttt ttaaagatga gatgaaacgg    480
cggcgggatt tctatgcagc gtacccccctg acggaggtcc ccaacggctc caatgaggac    540
cgcggcgagg tcttggagca ggagaagggg gctctgagtg atgatgagat tgtcagcctc    600
tccatcgaat tctacgaagg tgccagggac cgggatgaga agaagggccc cctggagaat    660
ggggatgggg acaaagagaa aacaggggtg cgcttcctgc gatgcccagc agccatgacc    720
gtcatgcatc ttgccaagtt tctccgcaac aagatggatg tgcccagcaa gtacaaggtg    780
gaggttctgt acgaggacga gccactgaag gaatactaca ccctcatgga catcgcctac    840
atctacccct ggcggcggaa cgggcctctc cccctcaagt accgtgtcca gccagcctgc    900
aagcggctca ccctagccac ggtgcccacc ccctccgagg gcaccaacac cagcgggggcg    960
tccgagtgtg agtcagtcag cgacaaggct cccagccctg ccaccctgcc agccacctcc   1020
tcctccctgc ccagcccagc caccccatcc catggctctc ccagttccca tgggcctcca   1080
gccacccacc ctacctcccc cactccccct tcgacagcca gtggggccac cacagctgcc   1140
aacgggggta gcttgaactg cctgcagaca ccatcctcca ccagcagggg gcgcaagatg   1200
actgtcaacg gcgctcccgt gcccccctta acttgaggcc agggaccctc tcccttcttc   1260
cagccaagcc tctccactcc ttccactttt tctgggccct tttttccact tcttctactt   1320
tccccagctc ttcccacctt gggggtgggg ggcgggtttt ataaataaat atatatatat   1380
atgtacatag gaaaaaccaa atatacatac ttattttcta tggaccaacc agattaattt   1440
aaatgccaca ggaaacaaac tttatgtgtg tgtgtatgtg tggaaaatgg tgttcatttt   1500
ttttgggggg ggtcttgtgt aatttgctgt tttgggggt gcctggagat gaactggatg   1560
ggccactgga gtctcaataa agctctgcac catcctcgct gtttcccaag gcaggtggtg   1620
tgttggggggc cccttcagac ccaaagcttt aggcatgatt ccaactggct gcatatagga   1680
gtcagttaga attgtttctt tctctccccg tttctctccc catcttggct gctgtcctgc   1740
ctctgaccag tggccgcccc ccgcgttgtt gaatgtccag aaattgctaa gaacagtgcc   1800
ttttacaaat gcagtttatc cctggttctg aggagcaagt gcagggtgga ggtggcacct   1860
gcatcacctc ctcctcttgc agtggaaact ttgtgcaaag aatagatagt tctgcctctt   1920
tttttttttt ttcctgtgtg tgtggccttt gcatcattta tcttgtggaa aagaagattc   1980
aggccctgag aggtctcagc tcttggagga gggctaaggc tttagcattg tgaagcgctg   2040
cacccccacc aaccttaccc tcaccgggga accctcacta gcaggactgg tggtggagtc   2100
tcacctgggg cctagagtgg aagtgggggt gggttaacct cacacaagca cagatcccag   2160
actttgccag aggcaaacag ggaattccgc cgatactgac gggctccagg agtcgtcgcc   2220
acactcg                                                             2227
```

<210> 55

<211> 4283

<212> DNA

<213> Homo sapiens

<400> 55

```
ttgcgggaaa gagccaaacc ctggcgttgg ggggcccggg cggggagccc ctcccgcggt     60
ccacagcgac gcctgcccag ccctcctccc cttccggctc cggcacgggg ccccgaggcg    120
```

```
ttcggaggcc aggcgggttt ctgtcaggcc cggggaggag gggcgggcgg ggcggccgct    180
gcctccccgg gacgggccgt accacgcgga cggggaggac ggggccaggg gactgcaggg    240
cggctgcacc gcccgggggc ggggtgcgga gcgggccggc gggctccccg gggcggggcg    300
ggagggcggg gcgtggggcg gacggaacca ccggggcggg gtgggaggta acgggacggg    360
cgcgaccatg gcgcggtgag ggagcggggg tggggatcgg tccgggggag gcctgaggcc    420
gctggcttgt gcgctgtctc cgccgccccc ctctttcgcc gccgccgccg ccgcccccggg   480
catgtcgtcc aactgcacca gcaccacggc ggtggcggtg gcgccgctca gcgccagcaa    540
gaccaagacc aagaagaagc atttcgtgtg ccagaaagtg aagctattcc gggccagcga    600
gccgatcctc agcgtcctga tgtggggggt gaaccacacg atcaatgagc tgagcaatgt    660
tcctgttcct gtcatgctaa tgccagatga cttcaaagcc tacagcaaga tcaaggtgga    720
caatcatctc ttcaataagg agaacctgcc cagccgcttt aagtttaagg agtattgccc    780
catggtgttc cgaaaccttc gggagaggtt tggaattgat gatcaggatt accagaattc    840
agtgacgcgc agcgccccca tcaacagtga cagccagggt cggtgtggca cgcgtttcct    900
caccacctac gaccggcgct ttgtcatcaa gactgtgtcc agcgaggacg tggcggagat    960
gcacaacatc ttaaagaaat accaccagtt tatagtggag tgtcatggca acacgctttt   1020
gccacagttc ctgggcatgt accgcctgac cgtggatggt gtggaaacct acatggtggt   1080
taccaggaac gtgttcagcc atcggctcac tgtgcatcgc aagtatgacc tcaagggttc   1140
tacggttgcc agagaagcga gcgacaagga gaaggccaag gacttgccaa cattcaaaga   1200
caatgacttc ctcaatgaag ggcagaagct gcatgtggga gaggagagta aaaagaactt   1260
cctggagaaa ctgaagcggg acgttgagtt cttggcacag ctgaagatca tggactacag   1320
cctgctggtg ggcatccacg acgtggaccg ggcagagcag gaggagatgg aggtggagga   1380
gcgggcagag gacgaggagt gtgagaatga tggggtgggt ggcaacctac tctgctccta   1440
tggcacacct ccggacagcc ctggcaacct cctcagcttt cctcggttct ttggtcctgg   1500
ggaattcgac ccctctgttg acgtctatgc catgaaaagc catgaaagtt cccccaagaa   1560
ggaggtgtat ttcatggcca tcattgatat cctcacgcca tacgatacaa agaagaaagc   1620
tgcacatgct gccaaaacgg tgaaacacgg ggcaggggcc gagatctcga ctgtgaaccc   1680
tgagcagtac tccaaacgct tcaacgagtt tatgtccaac atcctgacgt agttctcttc   1740
taccttcagc cagagccaga gagctggata tggggtcggg gatcgggagt tagggagaag   1800
ggtgtatttg ggctagatgg gagggtggga gcagagtcgg gtttgggagg gctttagcaa   1860
tgagactgca gcctgtgaca ccgaaagaga ctttagctga agaggagggg gatgtgctgt   1920
gtgtgcacct gctcacagga tgtaacccca ccttctgctt acccttgatt ttttctcccc   1980
atttgacacc caggttaaaa aggggttccc tttttggtac cttgtaacct tttaagatac   2040
cttggggcta gagatgactt cgtgggttta tttgggtttt gtttctgaaa tttcattgct   2100
ccaggtttgc tatttataat catatttcat cagcctaccc accctcccca tctttgctga   2160
gctctcagtt cccttcaatt aaagagatac ccagtagacc cagcacaagg gtccttccag   2220
aaccaagtgc tatggatgcc agattggaga ggtcagacac ctcgccctgc tgcatttgct   2280
cttgtctgga ttaactttgt aatttatgga gtattgtgca caacttcctc cacctttccc   2340
ttggattcaa gtgaaaactg ttgcattatt cctccatcct gtctggaata caccaggtca   2400
acaccagaga tctcagatca gaatcagaga tctcagaggg gaataagttc atcctcatgg   2460
gatggtgagg ggcaggaaag cggctgggct cttggacacc tggttctcag agaaccctgt   2520
gatgatcacc caagccccag gctgtcttag cccctggagt tcagaagtcc tctctgtaaa   2580
gcctgcctcc cactaggtca agaggaacta gagtaccttt ggatttatca ggaccctcat   2640
gtttaaatgg ttatttccct ttgggaaaac ttcagaaact gatgtatcaa atgaggccct   2700
gtgccctcga tctatttcct tcttccttct gacctcctcc caggcactct tacttctagc   2760
cgaactctta gctctgggca gatctccaag cgcctggagt gcttttagc agagacacct    2820
cgttaagctc cgggatgacc ttgtaggaga tctgtctccc tgtgcctgga gagttacagc   2880
cagcaaggtg cccccatctt agagtgtggt gtccaaacgt gaggtggctt cctagttaca   2940
tgaggatgtg atccaggaaa tccagtttgg aggcttgatg tgggttttga cctggcctca   3000
gccttggggc tgtttttcct tgttgccccg ctctagactt ttagcagatc tgcagcccac   3060
aggctttttt ggaaggagtg gcttcctgca ggtgttccac ctgccttcgg agcctgccac   3120
ccaggccctc agaactgagc cacaggctgc tctggccagg agagaaacag ctctgttgtt   3180
ctgcattggg ggaggtacat tcctgcatct tctcaccccc tcaaccagga actggggatt   3240
tgggatgaga tatggtcaga cttgtagata accccaaaga tgtgaagatc gcttgtgaaa   3300
ccattttgaa tgaatagatt ggtttcctgt ggctccctcc aaacctggcc aagcccagct   3360
tccgaagcag gaaccagcac tgtctctgtg cctgactcac agcatatagg tcaggaaaga   3420
atggagacgg cattcttgga cttcactggg gctgctggat tggatgggaa accttctgga   3480
agaggcagat gggggtcaaa ccactgcctt ggccccagga aggggccata ggtaggtctg   3540
aacaactgcc gcaagaccac tacatgactt agggaacttg aaaccaactg gctcatggag   3600
aaaacaaatt tgacttggga aagggattat gtaggaataa tgtttggact tgatttcccc   3660
acgtcataat gaagaatgga agtttggatc tgctcctcgt caggcgcagc atctctgaag   3720
cttggaaagc tgtcttccag cctccaaacc tggccaagcc cagcttccga agcaggaacc   3780
agcactgtct ctgtgcctga ctcacagcat ataggtcagg aaagaatgga gacggcattc   3840
ttggacttca ctggggctgc tggattggat gggaaacctt ctggaagagg cagatggggg   3900
tcaaaccact gccttggccc caggaagggg ccataggtag gtctgaacaa ctgccgcaag   3960
accactacat gacttaggga acttgaaacc aactggctca tggagaaaac aaatttgact   4020
tgggaaaggg attatgtagg aataatgttt ggacttgatt tccccacgtc ataatgaaga   4080
```

atggaagttt ggatctgctc ctcgtcaggc gcagcatctc tgaagcttgg aaagctgtct 4140
tccagcagcc tccgtggcct cgggttccta ccggcttctc tgcatttggt ctgctgatca 4200
tgttgccata atgtgtatgg aaagtgtaac acattcttac tggttaaaga cgactaccag 4260
gtatctaact tgtttaacat tga 4283

<210> 56

<211> 6140

<212> DNA

<213> Homo sapiens

<400> 56

gcggccgcag cctgagccag ggcccctcc ctcgtcagga ccggggcagc aagcaggccg 60
ggggcaggtc cgggcaccca ccatgcgagg cgagctctgg ctcctggtgc tggtgctcag 120
ggaggctgcc cgggcgctga gcccccagcc cggagcaggt cacgatgagg gcccaggctc 180
tggatgggct gccaaaggga ccgtgcgggg ctggaaccgg agagcccgag agagccctgg 240
gcatgtgtca gagccggaca ggacccagct gagccaggac ctgggtgggg gcaccctggc 300
catggacacg ctgccagata acaggaccag ggtggtggag gacaaccaca gctattatgt 360
gtcccgtctc tatggcccca gcgagcccca cagccgggaa ctgtgggtag atgtggccga 420
ggccaaccgg agccaagtga agatccacac aatactctcc aacacccacc ggcaggcttc 480
gagagtggtc ttgtcctttg atttcccttt ctacggcat cctctgcggc agatcaccat 540
agcaactgga ggcttcatct tcatggggga cgtgatccat cggatgctca cagctactca 600
gtatgtggcg cccctgatgg ccaacttcaa ccctggctac tccgacaact ccacagttgt 660
ttactttgac aatgggacag tctttgtggt tcagtgggac cacgtttatc tccaaggctg 720
ggaagacaag ggcagtttca ccttccaggc agctctgcac catgacggcc gcattgtctt 780
tgcctataaa gagatcccta tgtctgtccc ggaaatcagc tcctcccagc atcctgtcaa 840
aaccggccta tcggatgcct tcatgattct caatccatcc ccggatgtgc cagaatctcg 900
gcgaaggagc atctttgaat accaccgcat agagctggac cccagcaagg tcaccagcat 960
gtcggccgtg gagttcaccc cattgccgac ctgcctgcag catagggagct gtgacgcctg 1020
catgtcctca gacctgacct tcaactgcag ctggtgccat gtcctccaga gatgctccag 1080
tggctttgac cgctatcgcc aggagtggat ggactatggc tgtgcacagg aggcagaggg 1140
caggatgtgc gaggacttcc aggatgagga ccacgactca gcctcccctg acacttcctt 1200
cagcccctat gatggagacc tcaccactac ctcctcctcc ctcttcatcg acagcctcac 1260
cacagaagat gacaccaagt tgaatcccta tgcaggagga gacggccttc agaacaacct 1320
gtcccccaag acaaagggca ctcctgtgca cctgggcacc atcgtgggca tcgtgctggc 1380
agtcctcctc gtggcggcca tcatcctggc tggaatttac atcaatggcc accccacatc 1440
caatgctgcg ctcttcttca tcgagcgtag acctcaccac tggccagcca tgaagtttcg 1500
cagccacccc gaccattcca cctatgcgga ggtggagccc tcgggccatg agaaggaggg 1560
cttcatggag gctgagcagt gctgagaaca ccaagtctcc cctttgaaga ctttgaggcc 1620
acagaaaaga cagttaaagc aaagaagaga agtgactttt cctggcctct cccagcatgc 1680
cctgggctga gatgagatgg tggtttatgg ctccagagct gctgttcgct tcgtcagcac 1740
accccgaata ttgaagaggg ggccaaaaaa caaccacatg gatttttat aggaacaaca 1800
acctaatctc atcctgtttt gatgcaaggg ttctcttctg tgtcttgtaa ccatgaaaca 1860
gcagaagaac taacataact aactccattt ttgtttaagg ggcctttacc tattcctgca 1920
cctaggctag gataacttta gagcactgac ataaaacgca aaaacaggaa tcatgccgtt 1980
tgcaaaacta actctgggat taaaggggaa gcatgtaaac agctaactgt tttgttaaa 2040
gatttatagg aatgaggagg tttggctatt gtcacatgac agactgttag ccaaggacaa 2100
agaagttctg caaacctccc ctggaccctt gctggtgtcc agatgtctgc ggttgtcagc 2160
cccttccttt cccccgacct aaacataaaa gacaaggcaa agcccgcata atttaagac 2220
ggttctttag gacattagtc caccatcttc ttggtttgct ggctctccga aataaagtcc 2280
ctttccttgc tccaactcct tgtctctcaa cgtattggct atgacgcagc aagcagaatg 2340
aatttggact cagttacagg ctgtcaatgg tctgctctgt agcagtctca gagcctcccc 2400
gacccactac ctggagatag ccagatagcc agatgccctg ctcctggcca cctttaaagc 2460
ccctgcatat gacacaggtt aactaaagtc aagattgggg ctgctgcatt ccaggttccc 2520
tagactcaca agctggtcct tggccaggtg cagtggctca cgcctgtaat cccagcactt 2580
tgggaggctg aggcaggcgg atcacctgaa gtcagaagtt tgagaccagc ctggccaaca 2640
taattaaaat gtctctacta aaaatacaaa aaattagctg ggtgtggtga cgcttgcctg 2700
tatcccagct actcaggaag ctgagacacg agaatcactt gaacctggga ggcagaggtt 2760
gcagtgagct cagatagtgc cactgcactc cagcctgggt gacagagcga gactccgtct 2820
caaaaaaaaa aaaagaaagc agaacctcat ggctatagag ttggcatttt agccccagct 2880
tctgtagctc tgaaagccta aagaaggtat tctctccatc tgttaaacac agtatagtgg 2940
ctctcagccc ttggggcatg ttatcatggg agggaagtca aataagagga gagaaaagaa 3000
ctcaaggggg aaactgcatt tttaggcttt gctctcttac cttgcccttt ctactcagaa 3060

ccaataactt ctgcatcaaa acatgttaca gcctgcatca agggctttac cccaacctgc 3120
agcccagcct tccctgggtg agcttgctat gcgcagccac atttaccatg tggggctccc 3180
tattctgatg gcctgttcgg tgccgggttt actcactgcc ctgttctgat gtcagtgcct 3240
gtacatacct ccaaaggcag gacttgcctg ataaatattt ttcctcctct gaactggatt 3300
ttataggcat taaagacaag tcgggtggct agagggctcc ttgagacata cctagcaggg 3360
aactgcaggt ggattctgtt gagaggcaaa gcacctgagt ggttgggaca caggcagctg 3420
gcatgggagg gacttttttt gagacagggt ctcactgtgt cgcccagggc aaggatgccc 3480
aaagacacca ggttggagag gcacctgcca actacttgct ttccctggag cctgcatgtg 3540
cctgtggggt ggggaggcgt aggggtctac ggctgcctga gatgggtgtg cacagtgtgt 3600
gaagtaccta cctccttgcc ttgctggact gtcagccagt cgcagggccg gccacaagac 3660
ccatgtctcc atctggtcat actccatagc taccaagtta acctgctcta aactttggag 3720
aactggatct gtccaataaa cgcttatttg gccaagcctg atggctcgtg cctgtactcc 3780
cagcactttg ggaggctgag gtgggagggt tgcttgagcc caggggtttg agaccagctt 3840
gggcaacaac aacaaaaatg ccaggtgtgg tggggtgcac ctgtagtccc agctactagg 3900
gaggctgagc caggaggatc acttgagccc gggaggttga ggctgcagtg gggggtcata 3960
atcatgccac tgtactccag cctgggtgac agagtgagac cctgtctccg aaaaaaaaaa 4020
aaaaaaaaga acggaaaaag aaatgcttac attgtcaggg atcctgtaga caatcattaa 4080
ctctatgaga tgcttggttc tatttttttg ggagactttg tccaagtgtt ttggcttaag 4140
aaatccatag gcctctcttg gtgacacatc tctagtactt tttgtcataa acaaacaggc 4200
catctgccgc caaatacatc cactccccat gccactgaca tcctatgggt cagccaggct 4260
tgctttgact gaggccgagg catctggaac tttctctgcc tgcaggggct agcagcagag 4320
gcttcaccgc atcaccaccc cttcctccac tcctgacatt ctttcccttc agggatccaa 4380
aatggttggc cgagctccca gtgggaaaac gtgtgctaga gttggggagt gagatgagtg 4440
gtgctgtcca tggaatcagg ccacagcagg aactgcccca ctggccattt gagacacaca 4500
caggtggtaa atgctctgct ggtgggctgt gcttccctca ttcagagagc tctgttacag 4560
cccactgtgt cctttagaag cttgaaagga acccaactct ttgctgcact gtcctttttc 4620
ttcctcaaat tcagaccctc cttccaccgg caccccccta ctccaccctc agctcttcct 4680
tgcctggttt atcaagcaga gctgaggccc cacgtttcca actctgattg tcacttgcat 4740
cttcacaaag gataaaccac ggagcaactg gaaaaccatc agccaagcgt tcggatgagt 4800
ctggttattg gtccaccccc gaccagattc ccttacactt aactcacttc tttctttggc 4860
aatgaccctc atgacatgta taaatgggta tgactaagaa gaggctgtga tctaacattt 4920
atttgctgcc atttttttact ctggggagaa gcagccccaa ctcatcactg ggaaagaact 4980
ccccctgcaa accagctaaa tttgataatt taaacccccct gcccctaaaa cttctcacag 5040
agctggggag ttggtggcaa ctttccaagt caaggtcttg cttagaaagt ccttcactac 5100
atggccaggt gcagtggctc acgcctgtag tcccaggtac ttgggagcct gaggcaggag 5160
gattgcttga gctcaggagt tcaaggctgc agagagctat gatcatccca ctgcatttgt 5220
ttaaaaataa atttttaaaa tttgtgtgtt ttatcagggg tctcctgtac agtgtatctg 5280
tgtatgtttg tgtgtgtgtt tgtatacagc cttgtttaat gttttgagca ataagatatg 5340
cacacacagg tattttgttg ctaaagagat tggacaaggt tgtagctgtg ctcaggcttc 5400
agcttggttt gttaaattga gagataaaca atgacaagag ctgccagcca accacactat 5460
tcaaaaagca aagtgttcac cactaaagct aaccattcat ctggttgcag gcaaggctaa 5520
ggctctctct cctctagttc ctggaacaga ctcacagatt ggcatgaagc actgatcagg 5580
ggctgcactc agactccctg gccaagcaaa cctacaccag aagagtcagt gtcacagata 5640
tgatgcggcc aatctctgtc tccaaaaacc tacctgaact taatggtaga attcaaagat 5700
ctggggactg agggcaccca gccttctaaa acacaatgta ttcatgtgtt tagtgtaaac 5760
tctctgcatg gattctcagt gttaataata aaaggaagca ttcttttaca actcctgctg 5820
tgtgcaaaag aaagtgcaaa ggatttggag tggcattccg aagatcacca cacatacctt 5880
ggttctgatg gctgctgaac tccgacttct tcgctgagac atgactgtgg gaacagcctc 5940
cagctatctg ctcatcagag gtgctttcct caacctcctg caccacctcc aagagaaaca 6000
gcctaaaaag aaaccccagc tgtttactta tattggtctg taaatccctg gaagtaaacc 6060
ccatgcattt ttatctactg tctgaggaca tacaataaat ctgagaaagt ctatgctgtc 6120
aaaaaaaaaa aaaaaaaaaa 6140

<210> 57

<211> 2098

<212> DNA

<213> Homo sapiens

<400> 57

gcaggagcac gtggagaggc cgggtagcca cagcggcagc tccagcccgg cccggcagcg 60
acatggaaga tatacaaaca aatgcggaac tgaaaagcac tcaggagcag tctgtgcccg 120
cagaaagtgc agcggttttg aatgactaca gtttaaccaa atctcatgaa atggaaaatg 180

tggacagtgg agaaggccca gccaatgaag atgaagacat aggagatgat tcaatgaaag 240
tgaaagatga atacagtgaa agagatgaga atgttttaaa gtcagaaccc atgggaaatg 300
cagaagagcc tgaaatccct tacagctatt caagagaata taatgaatat gaaaacatta 360
agttggagag acatgttgtc tcattcgata gtagcaggcc aaccagtgga aagatgaact 420
gcgatgtgtg tggattatcc tgcatcagct tcaatgtctt aatggttcat aagcgaagcc 480
atactggtga acgcccattc cagtgtaatc agtgtggggc atcttttact cagaaaggta 540
acctcctccg ccacattaaa ctgcacacag ggaaaaaacc ttttaagtgt cacctctgca 600
actatgcatg ccaaagaaga gatgcgctca cggggcatct taggacacat tctgtggaga 660
aaccctacaa atgtgagttt tgtggaagga gttacaagca gagaagttcc cttgaggagc 720
acaaggagcg ctgccgtaca tttcttcaga gcactgaccc aggggacact gcaagtgcgg 780
aggcaagaca catcaaagca gagatgggaa gtgaaagagc tctcgtactg gacagattag 840
caagcaatgt ggcaaaacga aaaagctcaa tgcctcagaa attcattggt gagaagcgcc 900
actgctttga tgtcaactat aattcaagtt acatgtatga gaaagagagt gagctcatac 960
agacccgcat gatggaccaa gccatcaata acgccatcag ctatcttggc gccgaagccc 1020
tgtgcccctt ggtccagaca ccgcctgctc ccacctcgga gatggttcca gttatcagca 1080
gcatgtatcc catagccctc acccgggctg agatgtcaaa cggtgcccct caagagctgg 1140
aaaggaaaag catcctcctt ccagagaaga gcgtgccttc tgagagaggc ctctctccca 1200
acaatagtgg ccacgactcc acggacactg acagcaacca tgaagaacgc cagaatcaca 1260
tctatcagca aaatcacatg gtcctgtctc gggcccgcaa tgggatgcca cttctgaagg 1320
aggttccccg ctcttacgaa ctcctcaagc ccccgcccat ctgcccaaga gactctgtca 1380
aagtgatcga caaggaaggg gaggtgatgg atgtgtatcg gtgtgaccac tgccgcgtcc 1440
tcttcctgga ctatgtgatg ttcacgattc acatgggctg ccacggcttc cgtgaccctt 1500
tcgagtgtaa catgtgtgga gatcgaagcc atgatcggta tgaattctcg tctcacatag 1560
ccagaggaga acacagaagc ctgctgaagt gaatatctgg tctcagggat tgctcctatg 1620
tattcagcat cgtttctaaa aacagttgac ctcgcctaac agattgctct caaaacatac 1680
tcagttccaa acttcttttc ataccatttt tagctgtgtt cacaggggta gccagagaaa 1740
cactgtcttc cttcagaaat tattcgcagg tctagcatat tattactttt gtgaaacctt 1800
tgttttccca tcagggactt gaattttatg gaatttaaaa gccaaaaagg tatttggtca 1860
ttatcttcta cagcagtgga atgagtggtc ccggagatgt gctatatgaa acattctttc 1920
tgagatatat caaccacacg tggaaaagcc tttcagtcat acatgcaaat ccacaaagag 1980
gaagagctga ccagctgacc ttgctgggaa gcctcaccct tctgcccttc acaggctgaa 2040
gggttaagat ctaatctccc taatctaaat gacagtctaa gagtaagtaa aagaacag 2098

<210> 58

<211> 2947

<212> DNA

<213> Homo sapiens

<400> 58

atgccaattc ctcctccccc gccaccccca cctggtcctc ctccacctcc cacatttcat 60
caggcaaaca cagagcagcc caagctgagt agagatgagc agcggggtcg aggcgccctc 120
ttacaggaca tttgcaaagg gaccaagctg aagaaggtga ccaacattaa tgatcggagt 180
gctcccatcc tcgagaagcc gaaaggaagc agtggtggct atggctctgg aggagctgcc 240
ctgcagccca agggaggtct cttccaagga ggagtgctga agcttcgacc tgtgggagcc 300
aaggatggtt cagagaacct agctggtaag ccagccctgc aaatccccag ttctcgagct 360
gctgccccaa ggcctccagt atctgccgcc agcgggcgtc ctcaggatga tacagacagc 420
agccgggcct cactcccaga actgccccgg atgcagagac cctctttacc ggacctctct 480
cggcctaata ccaccagcag tacgggcatg aagcacagct cctctgcccc tcccccacca 540
cccccagggc ggcgtgccaa cgcacccccc acacctctgc ctatgcacag cagcaaagcc 600
cccgcctaca acagagagaa acccttgcca ccgacgcctg gacaaaggct tcaccctggt 660
cgagagggac ctcctgctcc acccccagtc aaaccacctc cttcccctgt gaatatcaga 720
acaggaccaa gtggccagtc tctggctcct cctcctccgc cttaccgcca gcctcctggg 780
gtccccaatg gaccctctag ccccactaat gagtcagccc ctgagctgcc acagagacac 840
aattctttgc ataggaagac accagggcct gtcagaggcc tagcacctcc tccacccacc 900
tcggcctccc catctttact gagtaatagg ccacctcccc cagcccgaga ccctcccagt 960
cggggagcag ctcctccacc cccaccacct gtgatccgaa atggtgccag ggatgctccc 1020
cctcccccac caccataccg aatgcatggg tcagaacccc cgagccgagg aaagccccca 1080
cctccacccct caaggacgcc agctgggcca ccccctcctc ctccaccgcc cctgaggaat 1140
ggccacagag attctatcac cactgtccgg tctttcttgg atgatttga gtcaaagtat 1200
tccttccatc cagtagaaga ctttcctgct ccagaagaat ataaacactt tcagaggata 1260
tatcccagca aaacaaaccg agctgcccgt ggagccccac ctctgccacc cattctcagg 1320
tgaagcctgg cttggtcccg ttcctcagga aaaggatgga ccttctcttc ttctcagatg 1380

```
gtcccttcca ttcccctgaa acctgcatga gagctcctaa catgtttctc caatgcaatc   1440
aagccctaga ctccaaatgt cctcccagct cacctccatc tatgcatctc atctctggat   1500
ttggtgatca gactctatat tgacagtagg atctcaaacc ctgcatccat ccttcctcca   1560
gcaagccctg ctagccacat gaggaacaag tttccgtgtc ttctgccttc ctcttgggga   1620
aaggtgcctt gttgtgatga attaactcac tgttagggca gggtggagaa tggtactcct   1680
tccttctcct gtccactgtg gggaagcttt ggcaggtata ttatatttca tcatttagga   1740
ggctggcatg accaggactt atgggtggga ggggagcatt tttagtgaag caagaaagga   1800
gtttgccaag aagtgatctg ttttaaaggt catatttgga gaaagggcaa ggaattgggt   1860
ctgctttatt tttgggggta ttttgttttt gttctcacct gctgcccccc caccccacca   1920
ccccagggat aaattggata taaacactaa atactaatca gttgaactta acatttaata   1980
aaaagaaagg gtgaaataaa ctgaagacca ttttagaact agtcagttct ctgcagcaaa   2040
gggaacagga gccatttgaa ccctctggga cccctcaccc cactgcttca gggtgctagg   2100
ctgagggatg tttttcctcc cccttaccgc ccatgccctt gaaagaaaag tcactttttg   2160
tggagggcat cattcattcc tgattcacaa accccaaaaa cctctggtgg gagataggaa   2220
gatagggcgt gggcctgggc cttaacctca atcttgtgtc tgcctcagtc ttttctgact   2280
ggccctgaag ttgtcagtgg ctctttctgt ccttcagccc ctggaaggtg ctccaggata   2340
acaaagaagg gcaggttgaa gcccctcatg gaaggagctg gctttgtggg gctgcaaagg   2400
acttttaagt cctgcctgta ctgaagttca cagcccacct gactgagcag actcttcctg   2460
ttcctttctc taccaccctt gccttcccag gactgcacgg tttaacacag cagagtacag   2520
aagggtgaag aagtgagcag aggcttatga agatattcag atactcttct atgccaggaa   2580
gcacaaagac tttgttgaga tttgcctcag ttcagtagat cttccttggc agccagccat   2640
aggttgtttc tttgtcttcc gggtcctaaa gagcacagag aaaatggagg tccccagtct   2700
aggtaggaag ctgattggat gaggacttct ttttttccga cagcaggatg gggctcttgg   2760
gctccacaca ccagatgctt tggttttcta caactgttgc tatgtgtaga gggtgctcag   2820
agcgtggcat gagagcaagg agaccatggc tactctttga aatggatggg gaaaattagc   2880
ttaaaaattt aatcacgaga ttgcgccact gcactccagc ctgggcgaca gagccagact   2940
ccgtctc                                                             2947
```

<210> 59

<211> 784

<212> DNA

<213> Homo sapiens

<400> 59

```
gagcggttgc gcagtgaagg ctagacccgg tttactggaa ttgctctggc gatcgagggg     60
tcctagtaca ccgcaatcat gtctattatg tcctataacg gaggggccgt catggccatg    120
aaggggaaga actgtgtggc catcgctgca gacaggcgct tcgggatcca ggcccagatg    180
gtgaccacgg acttccagaa gatctttccc atgggtgacc ggctgtacat cggtctggcc    240
gggctcgcca ctgacgtcca gacagttgcc cagcgcctca agttccggct gaacctgtat    300
gagttgaagg aaggtcggca gatcaaacct tataccctca tgagcatggt ggccaacctc    360
ttgtatgaga aacggtttgg cccttactac actgagccag tcattgccgg gttggacccg    420
aagaccttta agcccttcat ttgctctcta gacctcatcg gctgccccat ggtgactgat    480
gactttgtgg tcagtggcac ctgcgccgaa caaatgtacg gaatgtgtga gtccctctgg    540
gagcccaaca tggatccgga tcacctgttt gaaaccatct cccaagccat gctgaatgct    600
gtggaccggg atgcagtgtc aggcatggga gtcattgtcc acatcatcga gaaggacaaa    660
atcaccacca ggacactgaa ggcccgaatg gactaaccct gttcccagag cccacttttt    720
tttctttttt tgaaataaaa tagcctgtct ttcaaaaaaa aaaaaaaaaa aaaaaaaaaa    780
aaaa                                                                 784
```

<210> 60

<211> 3033

<212> DNA

<213> Homo sapiens

<400> 60

```
atactcctaa gctcctcccc cggcggcgag ccagggagaa aggatggccg gcctggcggc     60
gcggttggtc ctgctagctg ggcagcggc gctggcgagc ggctcccagg gcgaccgtga    120
gccggtgtac cgcgactgcg tactgcagtg cgaagagcag aactgctctg gggcgctct    180
```

```
gaatcacttc cgctcccgcc agccaatcta catgagtcta gcaggctgga cctgtcggga     240
cgactgtaag tatgagtgta tgtgggtcac cgttgggctc tacctccagg aaggtcacaa     300
agtgcctcag ttccatggca agtggccctt ctcccggttc ctgttctttc aagagccggc     360
atcggccgtg gcctcgtttc tcaatggcct ggccagcctg gtgatgctct gccgctaccg     420
caccttcgtg ccagcctcct cccccatgta ccacacctgt gtggccttcg cctgggtgtc     480
cctcaatgca tggttctggt ccacagtctt ccacaccagg gacactgacc tcacagagaa     540
aatggactac ttctgtgcct ccactgtcat cctacactca atctacctgt gctgcgtcag     600
gtgagcctgc ctgggtggct gcaggggcaa aatcgaaccc tgggggcaga aaggggtcac     660
ccagccttcc cctgggggcc ttcttcacta gtctcccaac acctacgccc cccaaccccc     720
aacacatcag ctgtcctggg tgaggactct ggggtaggac tgggggccct ggctcctgac     780
aaggagctgt agcacttgct gcccagctgt ggcctgtttg gtggggagag gggtagtgac     840
ttcaggggcc atgcaccaat gttgggggga ggagatgctt cagggaatgc tgctctgggg     900
atgggccacc tgccctctga gcaaccctgg acggtggggc aggaccgtgg ggctgcagca     960
cccagctgtg gtcagtgcct tccgggctct cctgctgctc atgctgaccg tgcacgtctc    1020
ctacctgagc ctcatccgct tcgactatgg ctacaacctg gtggccaacg tggctattgg    1080
cctggtcaac gtggtgtggt ggctggcctg gtgcctgtgg aaccagcggc ggctgcctca    1140
cgtgcgcaag tgcgtggtgg tggtcttgct gctgcagggg ctgtccctgc tcgagctgct    1200
tgacttccca ccgctcttct gggtcctgga tgcccatgcc atctggcaca tcagcaccat    1260
ccctgtccac gtcctctttt tcagctttct ggaagatgac agcctgtacc tgctgaagga    1320
atcagaggac aagttcaagc tggactgaag accttggagc gagtctgccc cagtggggat    1380
cctgcccccg ccctgctggc ctcccttctc ccctcaaccc ttgagatgat ttctctttt     1440
caacttcttg aacttggaca tgaaggatgt gggcccagaa tcatgtggcc agcccacccc    1500
ctgttggccc tcaccagcct tggagtctgt tctagggaag gcctcccagc atctgggact    1560
cgagagtggg cagcccctct acctcctgga gctgaactgg ggtggaactg agtgtgttct    1620
tagctctacc gggaggacag ctgcctgttt cctccccacc agcctcctcc ccacatcccc    1680
agctgcctgg ctgggtcctg aagccctctg tctacctggg agaccaggga ccacaggcct    1740
tagggataca gggggtcccc ttctgttacc accccccacc ctcctccagg acaccactag    1800
gtggtgctgg atgcttgttc tttggccagc caaggttcac ggcgattctc cccatgggat    1860
cttgagggac caagctgctg ggattgggaa ggagtttcac cctgaccgtt gccctagcca    1920
ggttcccagg aggcctcacc atactccctt tcagggccag ggctccagca agcccagggc    1980
aaggatcctg tgctgctgtc tggttgagag cctgccaccg tgtgtcggga gtgtgggcca    2040
ggctgagtgc ataggtgaca gggccgtgag catgggcctg ggtgtgtgtg agctcaggcc    2100
taggtgcgca gtgtggagac gggtgttgtc ggggaagagg tgtggcttca aagtgtgtgt    2160
gtgcaggggg tgggtgtgtt agcgtgggtt aggggaacgt gtgtgcgcgt gctggtgggc    2220
atgtgagatg agtgactgcc ggtgaatgtg tccacagttg agaggttgga gcaggatgag    2280
ggaatcctgt caccatcaat aatcacttgt ggagcgccag ctctgcccaa gacgccacct    2340
gggcggacag ccaggagctc tccatggcca ggctgcctgt gtgcatgttc cctgtctggt    2400
gccccttgc ccgcctcctg caaacctcac agggtcccca cacaacagtg ccctccagaa     2460
gcagcccctc ggaggcagag gaaggaaaat ggggatggct ggggctctct ccatcctcct    2520
tttctccttg ccttcgcatg gctggccttc ccctccaaaa cctccattcc cctgctgcca    2580
gccccttgc catagcctga ttttggggag gaggaagggg cgatttgagg gagaagggga      2640
gaaagcttat ggctgggtct ggtttcttcc cttcccagag ggtcttactg ttccaggtg      2700
gccccagggc aggcaggggc cacactatgc ctgcgccctg gtaaaggtga cccctgccat    2760
ttaccagcag ccctggcatg ttcctgcccc acaggaatag aatggaggga gctccagaaa    2820
ctttccatcc caaaggcagt ctccgtggtt gaagcagact ggatttttgc tctgcccctg    2880
acccttgtc cctctttgag ggaggggagc tatgctagga ctccaacctc agggactcgg     2940
gtggcctgcg ctagcttctt ttgatactga aaacttttaa ggtgggaggg tggcaaggga    3000
tgtgcttaat aaatcaattc caagcctcac ctg                                 3033
```

<210> 61

<211> 1174

<212> DNA

<213> Homo sapiens

<400> 61

```
aagctcctcc cccggcggcg agccagggag aaaggatggc cggcctggcg gcgcggttgg      60
tcctgctagc tgggcagcg gcgctggcga gcggctccca gggcgaccgt gagccggtgt      120
accgcgactg cgtactgcag tgcgaagagc agaactgctc tgggggcgct ctgaatcact    180
tccgctcccg ccagccaatc tacatgagtc tagcaggctg gacctgtcgg gacgactgta    240
agtatgagtg tatgtgggtc accgttgggc tctacctcca ggaaggtcac aaagtgcctc    300
agttccatgg caagtggccc ttctcccggt tcctgttctt tcaagagccg gcatcggccg    360
tggcctcgtt tctcaatggc ctggccagcc tggtgatgct ctgccgctac cgcaccttcg    420
```

tgccagcctc ctccccatg taccacacct gtgtggcctt cgcctgggtg tccctcaatg 480
catggttctg gtccacagtc ttccacacca gggacactga cctcacagag aaaatggact 540
acttctgtgc ctccactgtc atcctacact caatctacct gtgctgcgtc aggaccgtgg 600
ggctgcagca cccagctgtg gtcagtgcct tccgggctct cctgctgctc atgctgaccg 660
tgcacgtctc ctacctgagc ctcatccgct tcgactatgg ctacaacctg gtggccaacg 720
tggctattgg cctggtcaac gtggtgtggt ggctggcctg gtgcctgtgg aaccagcggc 780
ggctgcctca cgtgcgcaag tgcgtggtgg tggtcttgct gctgcagggg ctgtccctgc 840
tcgagctgct tgacttccca ccgctcttct gggtcctgga tgcccatgcc atctggcaca 900
tcagcaccat ccctgtccac gtcctctttt tcagctttct ggaagatgac agcctgtacc 960
tgctgaagga atcagaggac aagttcaagc tggttgaagc agactggatt tttgctctgc 1020
ccctgacccc ttgtccctct ttgagggagg ggagctatgc taggactcca acctcaggga 1080
ctcgggtggc ctgcgctagc ttcttttgat actgaaaact tttaaggtgg gagggtggca 1140
agggatgtgc ttaataaatc aattccaagc ctca 1174

<210> 62

<211> 3167

<212> DNA

<213> Homo sapiens

<400> 62
aagctcctcc cccggcggcg agccagggag aaaggatggc cggcctggcg gcgcggttgg 60
tcctgctagc tggggcagcg gcgctggcga gcggctccca gggcgaccgt gagccggtgt 120
accgcgactg cgtactgcag tgcgaagagc agaactgctc tgggggcgct ctgaatcact 180
tccgctcccg ccagccaatc tacatgagtc tagcaggctg gacctgtcgg gacgactgta 240
agtatgagtg tatgtgggtc accgttgggc tctacctcca ggaaggtcac aaagtgcctc 300
agttccatgg caagtggccc ttctcccggt tcctgttctt tcaagagccg gcatcggccg 360
tggcctcgtt tctcaatggc ctggccagcc tggtgatgct ctgccgctac cgcaccttcg 420
tgccagcctc ctccccatg taccacacct gtgtggcctt cgcctggatg agaaaactga 480
ggcacagcaa ggctaaataa cttgcccaag gacacacagg aaatgcagag ccaggaactg 540
aaccctggca gtctggctgt agggcttgca ttcttaatga taccactacc tcccaaatct 600
gaggaaaggg tgtccctcaa tgcatggttc tggtccacag tcttccacac cagggacact 660
gacctcacag agaaaatgga ctacttctgt gcctccactg tcatcctaca ctcaatctac 720
ctgtgctgcg tcaggtgagc ctgcctgggt ggctgcaggg gcaaaatcga accctggggg 780
cagaaagggg tcacccagcc ttcccctggg ggccttcttc actagtctcc caacacctac 840
gcccccaac ccccaacaca tcagctgtcc tgggtgagga ctctggggta ggactggggg 900
ccctggctcc tgacaaggag ctgtagcact tgctgcccag ctgtggcctg tttggtgggg 960
agaggggtag tgacttcagg ggccatgcac caatgttggg gggaggagat gcttcaggga 1020
atgctgctct ggggatgggc cacctgccct ctgagcaacc ctggacggtg ggcaggacc 1080
gtggggctgc agcacccagc tgtggtcagt gccttccggg ctctcctgct gctcatgctg 1140
accgtgcacg tctcctacct gagcctcatc cgcttcgact atggctacaa cctggtggcc 1200
aacgtggcta ttggcctggt caacgtggtg tggtggctgg cctggtgcct gtggaaccag 1260
cggcggctgc ctcacgtgcg caagtgcgtg gtggtggtct tgctgctgca ggggctgtcc 1320
ctgctcgagc tgcttgactt cccaccgctc ttctgggtcc tggatgccca tgccatctgg 1380
cacatcagca ccatccctgt ccacgtcctc tttttcagct ttctggaaga tgacagcctg 1440
tacctgctga aggaatcaga ggacaagttc aagctggact gaagaccttg gagcgagtct 1500
gccccagtgg ggatcctgcc cccgccctgc tggcctccct tctcccctca acccttgaga 1560
tgatttctc ttttcaactt cttgaacttg gacatgaagg atgtgggccc agaatcatgt 1620
ggccagccca cccctgttg gccctcacca gccttggagt ctgttctagg gaaggcctcc 1680
cagcatctgg gactcgagag tgggcagccc ctctacctcc tggagctgaa ctggggtgga 1740
actgagtgtg ttcttagctc taccgggagg acagctgcct gtttcctccc caccagcctc 1800
ctccccacat ccccagctgc ctggctgggt cctgaagccc tctgtctacc tgggagacca 1860
gggaccacag gccttaggga tacaggggt cccttctgt taccaccccc caccctcctc 1920
caggacacca ctaggtggtg ctggatgctt gttctttggc cagccaaggt tcacggcgat 1980
tctccccatg ggatcttgag ggaccaagct gctgggattg ggaaggagtt tcaccctgac 2040
cgttgcccta gccaggttcc caggaggcct caccatactc cctttcaggg ccagggctcc 2100
agcaagccca gggcaaggat cctgtgctgc tgtctggttg agagcctgcc accgtgtgtc 2160
gggagtgtgg gccaggctga gtgcataggt gacagggccg tgagcatggg cctgggtgtg 2220
tgtgagctca ggcctaggtg cgcagtgtgg agacgggtgt tgtcggggaa gaggtgtggc 2280
ttcaaagtgt gtgtgtgcag gggtgggtg tgttagcgtg ggttagggga acgtgtgtgc 2340
gcgtgctggt gggcatgtga gatgagtgac tgccggtgaa tgtgtccaca gttgagaggt 2400
tggagcagga tgagggaatc ctgtcaccat caataatcac ttgtggagcg ccagctctgc 2460
ccaagacgcc acctgggcgg acagccagga gctctccatg gccaggctgc ctgtgtgcat 2520

gttccctgtc tggtgcccct ttgcccgcct cctgcaaacc tcacagggtc cccacacaac 2580
agtgccctcc agaagcagcc cctcggaggc agaggaagga aaatggggat ggctggggct 2640
ctctccatcc tcctttctc cttgccttcg catggctggc cttcccctcc aaaacctcca 2700
ttcccctgct gccagcccct ttgccatagc ctgatttgg ggaggaggaa ggggcgattt 2760
gagggagaag gggagaaagc ttatggctgg gtctggtttc ttcccttccc agagggtctt 2820
actgttccag ggtggcccca gggcaggcag gggccacact atgcctgcgc cctggtaaag 2880
gtgacccctg ccatttacca gcagccctgg catgttcctg ccccacagga atagaatgga 2940
gggagctcca gaaactttcc atcccaaagg cagtctccgt ggttgaagca gactggattt 3000
ttgctctgcc cctgacccct tgtccctctt tgagggaggg gagctatgct aggactccaa 3060
cctcagggac tcgggtggcc tgcgctagct tcttttgata ctgaaaactt ttaaggtggg 3120
agggtggcaa gggatgtgct taataaatca attccaagcc tcacctg 3167

<210> 63

<211> 2733

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (2694)..(2694)

<223> n=a, c, g or t

<220>

<221> misc_feature

<222> (2724)..(2724)

<223> n=a, c, g or t

<400> 63

agggagaaag gatggccggc ctggcggcgc ggttggtcct gctagctggg gcagcggcgc 60
tggcgagcgg ctcccagggc gaccgtgagc cggtgtaccg cgactgcgta ctgcagtgcg 120
aagagcagaa ctgctctggg ggcgctctga atcacttccg ctcccgccag ccaatctaca 180
tgagtctagc aggctggacc tgtcgggacg actgtaagta tgagtgtatg tgggtcaccg 240
ttgggctcta cctccaggaa ggtcacaaag tgcctcagtt ccatggcaag tggcccttct 300
cccggttcct gttctttcaa gagccggcat cggccgtggc ctcgtttctc aatggcctgg 360
ccagcctggt gatgctctgc cgctaccgca ccttcgtgcc agcctcctcc cccatgtacc 420
acacctgtgt ggccttcgcc tgggtgtccc tcaatgcatg gttctggtcc acagtcttcc 480
acaccaggga cactgaccta cagagaaaat ggactacttc tgtgcctcct gtatcctaca 540
ctcaatctac ctgtgctgcg tcaggaccgt ggggctgcag cacccagctg tggtcaagtg 600
ccttccgggc tctcctgctg ctcatgctga ccgtgcacgt ctcctacctg agcctcatcc 660
gcttcgacta tggctacaac ctggtggcca acgtggctat tggcctggtc aacgtggtgt 720
ggtggctggc ctggtgcctg tggaaccagc ggcggctgcc tcacgtgcgc aagtgcgtgg 780
tggtggtctt gctgctgcag gggctgtccc tgctcgagct gcttgacttc ccaccgctct 840
tctgggtcct ggatgcccat gccatctggc acatcagcac catccctgtc cacgtcctct 900
ttttcagctt tctggaagat gacagcctgt acctgctgaa ggaatcagag gacaagttca 960
agctggactg agaccttgga gcgaagtctg ccccagtggg gatcctgccc ccgccctgct 1020
ggcctccctt ctcccctcaa cccttgagat gattttctct tttcaacttc ttgaacttgg 1080
acatgaagga tgtgggccca gaatcatgtg gccagcccac cccctgttgg ccctcaccag 1140
ccttggagtc tgttctaggg aaggcctccc agcatctggg actcgagagt gggcagcccc 1200
tctacctcct ggactgaact ggggtggaac tgagtgtgtt cttagctcta ccgggaggac 1260
agctgcctgt ttcctcccca ccagcctcct ccccacatcc ccagctgcct ggctgggtcc 1320
tgaagccctc tgtctacctg ggagaccagg gtaccacagg ccttagggat acagggggtc 1380
cccttctgtt accacccccc accctcctcc aggacaccac taggtggtgc tggatgcttg 1440
ttctttggcc agccaaggtt cacggcgatt ctccccatgg gatcttgagg gaccaagctg 1500
ctgggattgg gaaggagttt caccctgacc gttgccctag ccaggttccc aggaggcctc 1560

accatactcc ctttcagggc cagggctcca gcaagcccag ggcaaggatc ctgtgctgct 1620
gtctggttga gagcctgcca ccgtgtgtcg ggagtgtggg ccaggctgag tgcataggtg 1680
acagggccgt gagcatgggc ctgggtgtgt gtgagctcag gcctaggtgc gcagtgtgga 1740
gacgggtgtt gtcggggaag aggtgtggct tcaaagtgtg tgtgtgcagg gggtgggtgt 1800
gttagcgtgg gttaggggaa cgtgtgtgcg cgtgctggtg ggcatgtgag atgagtgact 1860
gccggtgaat gtgtccacag ttgagaggtt ggagcaggat gagggaatcc tgtcaccatc 1920
aataatcact tgtggagcgc cagctctgcc caagacgcca cctgggcgga cagccaggag 1980
ctctccatgg ccaggctgcc tgtgtgcatg ttccctgtct ggtgcccctt tgcccgcctc 2040
ctgcaaacct cacagggtcc ccacacaaca gtgccctcca gaagcagccc ctcggaggca 2100
gaggaaggaa aatggggatg gctggggctc tctccatcct ccttttctcc ttgccttcgc 2160
atggctggcc ttcccctcca aaacctccat tcccctgctg ccagcccctt tgccatagcc 2220
tgattttggg gaggaggaag ggcgatttgg agggagaagg ggagaaagct tatggctggg 2280
tctggtttct tcccttccca gagggtctta ctgttccagg gtggccccag gcagcagggc 2340
cacactatgc ctgcgccctg gtaaaggtga cccctgccat ttaccagcag ccctggcatg 2400
ttcctgcccc acaggaatag aatggaggga gctccagaaa ctttccatcc caaaggcagt 2460
ctccgtggtt gaagcagact ggatttttgc tctgcccctg accccttgtc cctctttgag 2520
ggagggggagc tatgctagga ctccaacctc agggactcgg gtggcctgcg ctagcttctt 2580
ttgatactga aaacttttaa ggtgggaggg tggcaaggga tgtgcttaag cggccgcgaa 2640
ttcaaaaagc ttctcgagag tacttctaga gcggccgcgg gcccatcgat tttnccaccc 2700
gggtggggta cccaggtaag tgtnccccat atc 2733

<210> 64

<211> 2546

<212> DNA

<213> Homo sapiens

<400> 64

aagctcctcc cccggcggcg agccagggag aaaggatggc cggcctggcg gcgcggttgg 60
tcctgctagc tggggcagcg gcgctggcga gcggctccca gggcgaccgt gagccggtgt 120
accgcgactg cgtactgcag tgcgaagagc agaactgctc tgggggcgct ctgaatcact 180
tccgctcccg ccagccaatc tacatgagtc tagcaggctg gacctgtcgg gacgactgta 240
agtatgagtg tatgtgggtc accgttgggc tctacctcca ggaaggtcac aaagtgcctc 300
agttccatgg caagtggccc ttctcccggt tcctgttctt tcaagagccg gcatcggccg 360
tggcctcgtt tctcaatggc ctggccagcc tggtgatgct ctgccgctac cgcaccttcg 420
tgccagcctc ctcccccatg taccacacct gtgtggcctt cgcctgggtg tccctcaatg 480
catggttctg gtccacagtc ttccacacca gggacactga cctcacagag aaaatggact 540
acttctgtgc ctccactgtc atcctacact caatctacct gtgctgcgtc aggcctggtc 600
aacgtggtgt ggtggctggc ctggtgcctg tggaaccagc ggcggctgcc tcacgtgcgc 660
aagtgcgtgg tggtggtctt gctgctgcag ggggctgtccc tgctcgagct gcttgacttc 720
ccaccgctct tctgggtcct ggatgcccat gccatctggc acatcagcac catccctgtc 780
cacgtcctct ttttcagctt tctggaagat gacagcctgt acctgctgaa ggaatcagag 840
gacaagttca agctggactg aagaccttgg agcgagtctg ccccagtggg gatcctgccc 900
ccgccctgct ggcctccctt ctcccctcaa cccttgagat gattttctct tttcaacttc 960
ttgaacttgg acatgaagga tgtgggccca gaatcatgtg gccagcccac cccctgttgg 1020
ccctcaccag ccttggagtc tgttctaggg aaggcctccc agcatctggg actcgagagt 1080
gggcagcccc tctacctcct ggagctgaac tggggtggaa ctgagtgtgt tcttagctct 1140
accgggagga cagctgcctg tttcctcccc accagcctcc tccccacatc cccagctgcc 1200
tggctgggtc ctgaagccct ctgtctacct gggagaccag ggaccacagg ccttagggat 1260
acaggggggtc cccttctgtt accacccccc accctcctcc aggacaccac taggtggtgc 1320
tggatgcttg ttctttggcc agccaaggtt cacggcgatt ctccccatgg gatcttgagg 1380
gaccaagctg ctgggattgg gaaggagttt caccctgacc gttgccctag ccaggttccc 1440
aggaggcctc accatactcc ctttcagggc cagggctcca gcaagcccag ggcaaggatc 1500
ctgtgctgct gtctggttga gagcctgcca ccgtgtgtcg ggagtgtggg ccaggctgag 1560
tgcataggtg acagggccgt gagcatgggc ctgggtgtgt gtgagctcag gcctaggtgc 1620
gcagtgtgga gacgggtgtt gtcggggaag aggtgtggct tcaaagtgtg tgtgtgcagg 1680
gggtgggtgt gttagcgtgg gttaggggaa cgtgtgtgcg cgtgctggtg ggcatgtgag 1740
atgagtgact gccggtgaat gtgtccacag ttgagaggtt ggagcaggat gagggaatcc 1800
tgtcaccatc aataatcact tgtggagcgc cagctctgcc caagacgcca cctgggcgga 1860
cagccaggag ctctccatgg ccaggctgcc tgtgtgcatg ttccctgtct ggtgcccctt 1920
tgcccgcctc ctgcaaacct cacagggtcc ccacacaaca gtgccctcca gaagcagccc 1980
ctcggaggca gaggaaggaa aatggggatg gctggggctc tctccatcct ccttttctcc 2040
ttgccttcgc atggctggcc ttcccctcca aaacctccat tcccctgctg ccagcccctt 2100

tgccatagcc tgattttggg gaggaggaag gggcgatttg agggagaagg ggagaaagct 2160
tatggctggg tctggtttct tcccttccca gagggtctta ctgttccagg gtggccccag 2220
ggcaggcagg ggccacacta tgcctgcgcc ctggtaaagg tgacccctgc catttaccag 2280
cagccctggc atgttcctgc cccacaggaa tagaatggag ggagctccag aaactttcca 2340
tcccaaaggc agtctccgtg gttgaagcag actggatttt tgctctgccc ctgacccctt 2400
gtccctcttt gagggagggg agctatgcta ggactccaac ctcagggact cgggtggcct 2460
gcgctagctt cttttgatac tgaaaacttt taaggtggga gggtggcaag ggatgtgctt 2520
aataaatcaa ttccaagcct cacctg 2546

<210> 65

<211> 2683

<212> DNA

<213> Homo sapiens

<400> 65

aagctcctcc cccggcggcg agccagggag aaaggatggc cggcctggcg gcgcggttgg 60
tcctgctagc tggggcagcg gcgctggcga gcggctccca gggcgaccgt gagccggtgt 120
accgcgactg cgtactgcag tgcgaagagc agaactgctc tgggggcgct ctgaatcact 180
tccgctcccg ccagccaatc tacatgagtc tagcaggctg gacctgtcgg gacgactgta 240
agtatgagtg tatgtgggtc accgttgggc tctacctcca ggaaggtcac aaagtgcctc 300
agttccatgg caagtggccc ttctcccggt tcctgttctt tcaagagccg gcatcggccg 360
tggcctcgtt tctcaatggc ctggccagcc tggtgatgct ctgccgctac cgcaccttcg 420
tgccagcctc ctcccccatg taccacacct gtgtggcctt cgcctgggtg tccctcaatg 480
catggttctg gtccacagtc ttccacacca gggacactga cctcacagag aaaatggact 540
acttctgtgc ctccactgtc atcctacact caatctacct gtgctgcgtc aggaccgtgg 600
ggctgcagca cccagctgtg gtcagtgcct tccgggctct cctgctgctc atgctgaccg 660
tgcacgtctc ctacctgagc ctcatccgct tcgactatgg ctacaacctg gtggccaacg 720
tggctattgg cctggtcaac gtggtgtggt ggctggcctg gtgcctgtgg aaccagcggc 780
ggctgcctca cgtgcgcaag tgcgtggtgg tggtcttgct gctgcagggg ctgtccctgc 840
tcgagctgct tgacttccca ccgctcttct gggtcctgga tgcccatgcc atctggcaca 900
tcagcaccat ccctgtccac gtcctctttt tcagctttct ggaagatgac agcctgtacc 960
tgctgaagga atcagaggac aagttcaagc tggactgaag accttggagc gagtctgccc 1020
cagtggggat cctgcccccg ccctgctggc ctcccttctc ccctcaaccc ttgagatgat 1080
tttctctttt caacttcttg aacttggaca tgaaggatgt gggcccagaa tcatgtggcc 1140
agcccacccc ctgttggccc tcaccagcct tggagtctgt tctagggaag gcctcccagc 1200
atctgggact cgagagtggg cagcccctct acctcctgga gctgaactgg ggtggaactg 1260
agtgtgttct tagctctacc gggaggacag ctgcctgttt cctccccacc agcctcctcc 1320
ccacatcccc agctgcctgg ctgggtcctg aagccctctg tctacctggg agaccaggga 1380
ccacaggcct tagggataca gggggtcccc ttctgttacc acccccaccc ctcctccagg 1440
acaccactag gtggtgctgg atgcttgttc tttggccagc caaggttcac ggcgattctc 1500
cccatgggat cttgagggac caagctgctg ggattgggaa ggagtttcac cctgaccgtt 1560
gccctagcca ggttcccagg aggcctcacc atactccctt tcagggccag ggctccagca 1620
agcccagggc aaggatcctg tgctgctgtc tggttgagag cctgccaccg tgtgtcggga 1680
gtgtgggcca ggctgagtgc ataggtgaca gggccgtgag catgggcctg ggtgtgtgtg 1740
agctcaggcc taggtgcgca gtgtggagac gggtgttgtc ggggaagagg tgtggcttca 1800
aagtgtgtgt gtgcaggggg tgggtgtgtt agcgtgggtt aggggaacgt gtgtgcgcgt 1860
gctggtgggc atgtgagatg agtgactgcc ggtgaatgtg tccacagttg agaggttgga 1920
gcaggatgag ggaatcctgt caccatcaat aatcacttgt ggagcgccag ctctgcccaa 1980
gacgccacct gggcggacag ccaggagctc tccatggcca ggctgcctgt gtgcatgttc 2040
cctgtctggt gcccctttgc ccgcctcctg caaacctcac agggtcccca cacaacagtg 2100
ccctccagaa gcagcccctc ggaggcagag gaaggaaaat ggggatggct ggggctctct 2160
ccatcctcct tttctccttg ccttcgcatg gctggccttc ccctccaaaa cctccattcc 2220
cctgctgcca gcccctttgc catagcctga ttttggggag gaggaagggg cgatttgagg 2280
gagaagggga gaaagcttat ggctgggtct ggtttcttcc cttcccagag ggtcttactg 2340
ttccagggtg gccccagggc aggcaggggc cacactatgc ctgcgccctg gtaaaggtga 2400
cccctgccat ttaccagcag ccctggcatg ttcctgcccc acaggaatag aatggaggga 2460
gctccagaaa ctttccatcc caaaggcagt ctccgtggtt gaagcagact ggatttttgc 2520
tctgcccctg acccttgtc cctctttgag ggaggggagc tatgctagga ctccaacctc 2580
agggactcgg gtggcctgcg ctagcttctt ttgatactga aaacttttaa ggtgggaggg 2640
tggcaaggga tgtgcttaat aaatcaattc caagcctcac ctg 2683

<210> 66

<211> 2341

<212> DNA

<213> Homo sapiens

<400> 66
aagctcctcc cccggcggcg agccagggag aaaggatggc cggcctggcg gcgcggttgg 60
tcctgctagc tggggcagcg gcgctggcga gcggctccca gggcgaccgt gagccggtgt 120
accgcgactg cgtactgcag tgcgaagagc agaactgctc tgggggcgct ctgaatcact 180
tccgctcccg ccagccaatc tacatgagtc tagcaggctg gacctgtcgg gacgactgta 240
agtatgagtg tatgtgggtc accgttgggc tctacctcca ggaaggtcac aaagtgcctc 300
agttccatgg caagtggccc ttctcccggt tcctgttctt tcaagagccg gcatcggccg 360
tggcctcgtt tctcaatggc ctggccagcc tggtgatgct ctgccgctac cgcaccttcg 420
tgccagcctc ctcccccatg taccacacct gtgtggcctt cgcctgggtg tccctcaatg 480
catggttctg gtccacagtc ttccacacca gggacactga cctcacagag aaaatggact 540
acttctgtgc ctccactgtc atcctacact caatctacct gtgctgcgtc agctttctgg 600
aagatgacag cctgtacctg ctgaaggaat cagaggacaa gttcaagctg gactgaagac 660
cttggagcga gtctgcccca gtggggatcc tgcccccgcc ctgctggcct cccttctccc 720
ctcaaccctt gagatgattt tctcttttca acttcttgaa cttggacatg aaggatgtgg 780
gcccagaatc atgtggccag cccacccccct gttggccctc accagccttg gagtctgttc 840
tagggaaggc ctcccagcat ctgggactcg agagtgggca gcccctctac ctcctggagc 900
tgaactgggg tggaactgag tgtgttctta gctctaccgg gaggacagct gcctgtttcc 960
tccccaccag cctcctcccc acatccccag ctgcctggct gggtcctgaa gccctctgtc 1020
tacctgggag accagggacc acaggcctta gggatacagg gggtcccctt ctgttaccac 1080
cccccaccct cctccaggac accactaggt ggtgctggat gcttgttctt tggccagcca 1140
aggttcacgg cgattctccc catggatctt tgagggacca agctgctggg attgggaagg 1200
agtttcaccc tgaccgttgc cctagccagg ttcccaggag gcctcaccat actccctttc 1260
agggccaggg ctccagcaag cccagggcaa ggatcctgtg ctgctgtctg gttgagagcc 1320
tgccaccgtg tgtcgggagt gtgggccagg ctgagtgcat aggtgacagg gccgtgagca 1380
tgggcctggg tgtgtgtgag ctcaggccta ggtgcgcagt gtggagacgg gtgttgtcgg 1440
ggaagaggtg tggcttcaaa gtgtgtgtgt gcagggggtg ggtgtgttag cgtgggttag 1500
gggaacgtgt gtgcgcgtgc tggtgggcat gtgagatgag tgactgccgg tgaatgtgtc 1560
cacagttgag aggttggagc aggatgaggg aatcctgtca ccatcaataa tcacttgtgg 1620
agcgccagct ctgcccaaga cgccacctgg gcggacagcc aggagctctc catggccagg 1680
ctgcctgtgt gcatgttccc tgtctggtgc ccctttgccc gcctcctgca aacctcacag 1740
ggtccccaca caacagtgcc ctccagaagc agcccctcgg aggcagagga aggaaaatgg 1800
ggatggctgg ggctctctcc atcctccttt tctccttgcc ttcgcatggc tggccttccc 1860
ctccaaaacc tccattcccc tgctgccagc ccctttgcca tagcctgatt ttggggagga 1920
ggaaggggcg atttgaggga gaaggggaga aagcttatgg ctgggtctgg tttcttccct 1980
tcccagaggg tcttactgtt ccagggtggc cccagggcag gcaggggcca cactatgcct 2040
gcgccctggt aaaggtgacc cctgccattt accagcagcc ctggcatgtt cctgccccac 2100
aggaatagaa tggagggagc tccagaaact ttccatccca aaggcagtct ccgtggttga 2160
agcagactgg atttttgctc tgcccctgac cccttgtccc tctttgaggg aggggagcta 2220
tgctaggact ccaacctcag ggactcgggt ggcctgcgct agcttctttt gatactgaaa 2280
acttttaagg tgggagggtg gcaagggatg tgcttaataa atcaattcca agcctcacct 2340
g 2341

<210> 67

<211> 2109

<212> DNA

<213> Homo sapiens

<400> 67
gattcggccg gagctgccag cggggaggct gcagccgcgg gttgttacag ctgctggagc 60
agcagcggcc cccgctcccg ggaaccgttc ccgggccgtt gatcttcggc cccacacgaa 120
cagcagagag gggcagcagg atgaatgtgg gcacagcgca cagcgaggtg aaccccaaca 180
cgcgggtgat gaacagccgt ggcatctggc tctcctacgt gctggccatc ggtctcctcc 240
acatcgtgct gctgagcatc ccgtttgtga gtgtccctgt cgtctggacc ctcaccaacc 300
tcattcacaa catgggcatg tatatcttcc tgcacacggt gaaggggaca ccctttgaga 360

```
ccccggacca gggcaaggcg aggctgctaa cccactggga gcagatggat tatggggtcc     420
agttcacggc ctctcggaag ttcttgacca tcacacccat cgtgctgtac ttcctcacca     480
gcttctacac taagtacgac cagatccatt ttgtgctcaa caccgtgtcc ctgatgagcg     540
tgcttatccc caagctgccc cagctccacg gagtccggat ttttggaatc aataagtact     600
gagagtgcag ccccttcccc tgcccagggt ggcaggggag gggtagggta aaaggcatgt     660
gctgcaacac tgaagacaga aagaagaagc ctctggacac tgccagagat gggggttgag     720
cctctggcct aatttccccc ctcgcttccc ccagtagcca acttggagta gcttgtagtg     780
gggttggggt aggccccctg ggctctgacc ttttctgaat tttttgatct cttccttttg     840
ctttttgaat agagactcca tggagttggt catggaatgg gctgggctcc tgggctgaac     900
atggaccacg cagttgcgac aggaggccag gggaaaaacc cctgctcact tgtttgccct     960
caggcagcca aagcacttta acccctgcat agggagcaga gggcggtacg gcttctggat    1020
tgtttcactg tgattcctag gtttttttcga tgccatgcag tgtgtgcttt tgtgtatgga    1080
agcaagtgtg ggatgggtct ttgcctttct gggtagggag ctgtctaatc caagtcccag    1140
gcttttggca gcttctctgc aacccaccgt gggtcctggt tgggagtggg gagggtcagg    1200
ttggggaaag atggggtaga gtgtagatgg cttggttcca gaggtgaggg ggccagggct    1260
gctgccatcc tggcctggtg gaggttgggg agctgtagga gagctagtga gtcgagactt    1320
agaagaatgg ggccacatag cagcagagga ctggtgtaag ggagggaggg gtagggacag    1380
aagctagacc caatctcctt tgggatgtgg gcagggaggg aagcaggctt ggagggttaa    1440
tttacccaca gaatgtgata gtaatagggg agggaggctg ctgtgggttt aactcctggg    1500
ttggctgttg ggtagacagg tggggaaaag gcccgtgagt cattgtaagc acaggtccaa    1560
cttggccctg actcctgcgg gggtatgggg aagctgtgac agaaacgatg ggtgctgtgg    1620
tcctctgcag gccctcaccc cttaacttcc tcatgcagac tggcactggg cagggcctct    1680
catgtggcag ccacatgtgg cgttgtgagg ccaccccatg tggggtctgt ggtgagagtc    1740
ctgtaggatc cctgctcaag cagcacagag gaaggggcaa gacgtggcct gtaggcactg    1800
tctcagcctg cagagaagaa agtgaggccg ggagcctgag cctgggctgg agccttctcc    1860
cctccccagt tggactaggg gcagtgttaa ttttgaaaag gtgtgggtcc ctgtgtcctt    1920
ttccaggggt ccaagggaac aggagaggtc actgggcctg ttttctccct cctgaccctg    1980
catctcccac cctgtgtatc atagggaact ttcaccttaa aatctttcta agcaaagtgt    2040
gaataggatt tttactccct ttgtacagta ttctgaggaa cgcaaataaa agggcaacat    2100
gtttctgtt                                                            2109
```

<210> 68

<211> 2423

<212> DNA

<213> Homo sapiens

<400> 68

```
gagagccgag ctagcgacga gcagtcgttg cggccgccgg cgccgcggga ggtggtggag      60
gcctagccgg agccgagagg tctcttgttc ccgtcccacg gtcccggcgt cacccctccg     120
gcgcccagtc cccgtcccgg aactcccggg cctgtcctgg gcccccggtc tgtgcactcc     180
gctcgccgca gcgcccggcc cgggccgcac ccgccggccc catgaggagg gacgtgaacg     240
gagtgaccaa gagcaggttt gagatgttct caaatagtga tgaagctgta atcaataaaa     300
aacttcccaa agaactcctg ttacggatat ttctcttttct agatgttgtt accctgtgcc     360
gctgtgctca ggtctccagg gcctggaatg ttctggctct ggatggcagt aactggcagc     420
gaattgacct atttgatttc cagagggata ttgagggccg agtagtggag aatatttcaa     480
aacgatgtgg gggctttttta cgaaagttaa gtcttcgtgg atgtcttgga gtgggagaca     540
atgcattaag aaccttttgca caaaactgca ggaacattga agtactgaat ctaaatgggt     600
gtacaaagac aacagacgct acatgtacta gccttagcaa gttctgttcc aaactcaggc     660
accttgactt ggcttcctgt acatcaataa caaacatgtc tctaaaagct ctgagtgagg     720
gatgtccact gttggagcag ttgaacattt cctggtgtga ccaagtaacc aaggatggca     780
ttcaagcact agtgaggggc tgtgggggtc tcaaggcctt attcttaaaa ggctgcacgc     840
agctagaaga tgaagctctc aagtacatag gtgcacactg ccctgaactg gtgactttga     900
acttgcagac ttgcttgcaa atcacagatg aaggtctcat tactatatgc agagggtgcc     960
ataagttaca atccctttgt gcctctggct gctccaacat cacagatgcc atcctgaatg    1020
ctctaggtca gaactgccca cggcttagaa tattggaagt ggcaagatgt tctcaattaa    1080
cagatgtggg ctttaccact ctagccagga attgccatga acttgaaaag atggacctgg    1140
aagagtgtgt tcagataaca gatagcacat taatccaact ttctatacac tgtcctcgac    1200
ttcaagtatt gagtctgtct cactgtgagc tgatcacaga tgatggaatt cgtcacctgg    1260
ggaatggggc ctgcgcccat gaccagctgg aggtgattga gctggacaac tgcccactaa    1320
tcacagatgc atccctggag cacttgaaga gctgtcatag ccttgagcgg atagaactct    1380
atgactgcca gcaaatcaca cgggctggaa tcaagagact caggacccat ttacccaata    1440
ttaaagtcca cgcctacttc gcacctgtca ctccaccccc atcagtaggg ggcagcagac    1500
```

agcgcttctg cagatgctgc atcatcctat gacaatggag gtggtcaacc ttggcgaact 1560
gagtatttaa tgacacttct agagctaccg tggagtctct ccagtggaag caacccccagt 1620
gttctgagca agggttacaa agtgagggag ggcagtgtcc agatccccag agccacacat 1680
acatacacat acacaccctt accccccatcc actctagctt tgtgaccatg ggactgaagt 1740
ttgtgatggc ttttttatca agtagattgg taaaatttaa ccattcctgt tgaggtgccc 1800
ataagaaaat cataggccaa gatagggagg ggcattccag caaacccccgt gttaatgcta 1860
ctgtggtttt taaatttttg tctaggggtt tctttgggga ttttagaaca gcatctgctg 1920
tcctccgggg tcaagaaaag catggaaaga caatatatga tgtacccagg gaccagaaag 1980
aaaatttctt tgcatcttag aaatggtaga cattcattgt gactaaagag cttctatgct 2040
tccttgtttc catgccaaca tgctgagcat gctcacaaag aaggctcgtc cattcctcct 2100
gtgttttagt atttggccca gaggtttcct aaatggttgc cttgaaatca ctgtggtcca 2160
aatgtaattc ttacacactc aaattatcac tgtctgtagc acacttgtgc acctgtctta 2220
cattctctgt tgctcccccc cacactcttg ctcagtctgt cacctgttca gtctgcttac 2280
tcactcaatt gttacccttt tgctgttgtc gtgtttacag tttgcatttt gaatgattag 2340
ttgggattac caaacatttt ttaaaaagat attatcaata aatatttttt taattctaaa 2400
ttttaaaaaa aaaaaaaaaa aaa 2423

<210> 69

<211> 1841

<212> DNA

<213> Homo sapiens

<400> 69

agctgggacc ggagggtgag cccggcagag gcagagacac acgcggagag gaggagaggc 60
tgagggaggg aggtggagaa ggacgggaga ggcagagaga ggagacacgc agagacactc 120
aggaggggag agacaccgag acgcagagac actcaggagg ggagagacac cgagacgcag 180
agacacccag gccggggagc gcgagggagc gaggcacaga cctggctcag cgagcgcggg 240
gggcgagccc cgagtcccga gagcctgggg gcgcgcccag cccgggcgcc gaccctcctc 300
ccgctcccgc gccctcccct cggcgggcac ggtatttta tccgtgcgcg aacagccctc 360
ctcctcctct cgccgcacag cccgccgcct gcgcgggggga gcccagcaca gaccgccgcc 420
gggaccccga gtcgcgcacc ccagccccac cgcccacccc gcgcgccatg gaccccaagg 480
accgcaagaa gatccagttc tcggtgcccg cgccccctag ccagctcgac ccccgccagg 540
tggagatgat ccggcgcagg agaccaacgc ctgccatgct gttccggctc tcagagcact 600
cctcaccaga ggaggaagcc tccccccacc agagagcctc aggagagggg caccatctca 660
agtcgaagag acccaacccc tgtgcctaca caccaccttc gctgaaagct gtgcagcgca 720
ttgctgagtc tcacctgcag tctatcagca atttgaatga gaaccaggcc tcagaggagg 780
aggatgagct gggggagctt cgggagctgg gttatccaag agaggaagat gaggaggaag 840
aggaggatga tgaagaagag gaagaagaag aggacagcca ggctgaagtc ctgaaggtca 900
tcaggcagtc tgctgggcaa aagacaacct gtggccaggg tctggaaggg ccctgggagc 960
gcccaccccc tctggatgag tccgagagag atggaggctc tgaggaccaa gtggaagacc 1020
cagcactaag tgagcctggg gaggaacctc agcgcccttc cccctctgag cctggcacat 1080
aggcacccag cctgcatctc ccaggaggaa gtggagggga catcgctgtt ccccagaaac 1140
ccactctatc ctcaccctgt tttgtgctct tcccctcgcc tgctagggct gcggcttctg 1200
acttctagaa gactaaggct ggtctgtgtt tgcttgtttg cccacctttg gctgataccc 1260
agagaacctg ggcacttgct gcctgatgcc cacccctgcc agtcattcct ccattcaccc 1320
agcgggaggt gggatgtgag acagcccaca ttggaaaatc cagaaaaccg ggaacaggga 1380
tttgcccttc acaattctac tccccagatc ctctcccctg gacacaggag acccacaggg 1440
caggacccta agatctgggg aaaggaggtc ctgagaacct tgaggtaccc ttagatcctt 1500
ttctacccac tttcctatgg aggattccaa gtcaccactt ctctcaccgg cttctaccag 1560
ggtccaggac taaggcgttt ttctccatag cctcaacatt ttgggaatct tcccttaatc 1620
acccttgctc ctcctgggtg cctggaagat ggactggcag agacctcttt gttgcgtttt 1680
gtgctttgat gccaggaatg ccgcctagtt tatgtccccg gtggggcaca cagcgggggg 1740
cgccaggttt tccttgtccc ccagctgctc tgcccctttc cccttcttcc ctgactccag 1800
gcctgaaccc ctcccgtgct gtaataaatc tttgtaaata a 1841

<210> 70

<211> 748

<212> DNA

<213> Homo sapiens

<400> 70
ggccgcgatg agcggggagc cggggcagac gtccgtagcg ccccctcccg aggaggtcga 60
gccgggcagt ggggtccgca tcgtggtgga gtactgtgaa ccctgcggct tcgaggcgac 120
ctacctggag ctggccagtg ctgtgaagga gcagtatccg ggcatcgaga tcgagtcgcg 180
cctcgggggc acaggtgcct ttgagataga gataaatgga cagctggtgt tctccaagct 240
ggagaatggg ggctttccct atgagaaaga tctcattgag gccatccgaa gagccagtaa 300
tggagaaacc ctagaaaaga tcaccaacag ccgtcctccc tgcgtcatcc tgtgactgca 360
caggactctg ggttcctgct ctgttctggg gtccaaacct tggtctccct ttggtcctgc 420
tgggagctcc ccctgcctct ttcccctact tagctcctta gcaaagagac cctggcctcc 480
actttgccct ttgggtacaa agaaggaata gaagattccg tggccttggg ggcaggagag 540
agacactctc catgaacact tctccagcca cctcataccc ccttcccagg gtaagtgccc 600
acgaaagccc agtccactct tcgcctcggt aatacctgtc tgatgccaca gattttattt 660
attctcccct aacccagggc aatgtcagct attggcagta aagtggcgct acaaacacta 720
aaaaaaaaaa aaaaaaaaaa aaaaaaa 748
<210> 71

<211> 795

<212> DNA

<213> Homo sapiens

<400> 71
tacggctgcg agaagacgac agaagctaga cccaatctcc tttgggatgt gggcagggag 60
ggaagcaggc ttggagggtt aatttaccca cagaatgtga tagtaatagg ggagggaggc 120
tgctgcgggt ttaactcctg ggttggctgt tgggtagaca ggtggggaaa aggcccgtga 180
gtcattgtaa gcacaggtcc aacttggccc tgactcctgc gggggtatgg ggaagctgtg 240
acagaaacga tgggtgctgt ggtcctctgc aggccctcac cccttaactt cctcatacag 300
actggcactg ggcagggcct ctcatgtggc agccacatgt ggcgttgtga ggccacccca 360
tgtggggtct gtggtgagag tcctgtagga tccctgctca agcagcacag aggaaggggc 420
aagacgtggc ctgtaggcac tgtctcagcc tgcagagaag aaagtgaggc cgggagcctg 480
agcctgggct ggagccttct cccctcccca gttggactag gggcagtgtt aatttgaaa 540
aggtgtgggt ccctgtgtcc tcttccaggg gtccaaggga acaggagagg tcactggcc 600
tgttttctcc ctcctgaccc tgcatctccc accccgtgta tcatagggaa ctttcacctt 660
aaaatctttc taagcaaagt gtgaatagga tttttactcc ctttgtacag tattctgaga 720
aacgcaaata aaagggcaac atgtttctgt taaaaaaaaa aaaaagtacg caaaaaaaaa 780
aaaaaaaaaa aaaaa 795
<210> 72

<211> 2356

<212> DNA

<213> Homo sapiens

<400> 72
ggcacgaggc cggaagtgac ctctagagcg gtggtgaaac tggcagttga cggctcctgg 60
gactagatcc cgcgaggtag cccccgaact atttctctac gttttctctt gatcctcccg 120
aaatcttcca gatccgcgta gtgaggaatc gtctccaccg tcatgggggg cggagacctg 180
aatctgaaga agagctggca cccgcagacc ctcaggaatg tggagaaagt gtggaaggcc 240
gagcagaagc atgaggctga gcggaagaag attgaggagc ttcagcggga gctgcgagaa 300
gagagagccc gggaagagat gcagcgctat gcggaggatg ttggggccgt caagaaaaaa 360
gaagaaaagt tggactggat gtaccagggt cctggtggga tggtgaaccg tgacgagtac 420
ctgctggggc gccccattga caaatatgtt tttgagaaga tggaggagaa ggaggcaggc 480
tgctcttctg aaacaggact tctcccaggc tctatctttg ccccatcagg tgccaattcc 540
cttcttgaca tggccagcaa gatccgggag gacccactct tcatcatcag gaagaaggag 600
gaggagaaaa aacgagaggt attaaataat ccagtgaaaa tgaagaaaat caaagaattg 660
ttgcaaatga gtctggaaaa aaaggagaag aagaaaaaga aggagaagaa aaagaagcac 720
aagaaacata agcacagaag ctcgagtagt gatcgttcca gcagcgagga tgagcacagt 780
gcagggagat cacagaagaa gatggcaaat tcctcccctg ttttgtccaa agtccctgga 840

```
tatggcttac aggtccggaa ctctgaccgt aaccagggtc ttcagggtcc tctgacagca     900
gagcaaaaga gagggcatgg gatgaagaac cattccagat ccagaagctc ctcccactca     960
cccccaagac atgccagcaa gaagagcacc agggaagcag ggtcccggga caggaggtct    1020
cgatccctgg gcagaaggtc acggtcccca agacccagca aactgcacaa ctctaaggtg    1080
aacaggagag agacaggcca aactaggagc ccatcaccta aaaaagaggt ctaccaaagg    1140
cgacatgctc ccggatacac cagaaaactc tctgcagagg aattagagcg aaaacggcaa    1200
gagatgatgg aaaacgccaa atggagggag gaggagagac tgaacatcct caagaggcat    1260
gctaaggatg aggaacggga gcagaggcta gagaagctgg actcccggga tgggaagttc    1320
atccaccgca tgaagctgga gagtgcatct acttcctccc tggaggatcg ggtgaagcgg    1380
aatatctact ctttacagag aacttcggta gctctggaga agaactttat gaaaagatga    1440
aaactgtccc ctctcttatt ggtttttcctg catttttccag ggaagctgct gaccccttaa    1500
ttctctttat aagagttcaa atgacttctt tcacagatgt caaaccacca gtgttcaaag    1560
tgaccctgct tcattgagtc ctgaaacagc tcacttcctt tgagagctag tgtgacttgc    1620
tttgtgggac actcagtaac tttgggtttt gactctttaa cgggtgggca ctggaccatc    1680
tcggtgggag tgcttgtgcc actctggaag gctgttccct ggggttgtga tgtttatcat    1740
gccacttcct tcttacctgt gccaacagac ctatttcact gcctcagcgt acaccagacc    1800
cttcagaaac ctctctggtg tcacccagat agattgtgct tactgagaca aatgaacgtt    1860
tacttgattt agaagataat gtgacagaat gatgtcaggt taggtcaaag ccaagggagt    1920
gacagaatct ggaaaatcaa acaatacaaa aagccctaaa tgaactgtta actatttgat    1980
ctttggatgt aaaattgtaa tgcgtatatg tacaaatgta caatttttac atgcttttaa    2040
aaaaggttag ctttgtgaaa ataccttgtt tggtcaatga ctttactggg taatagaacc    2100
acattgaacc ttgatggcaa gtaatacaat aaggcaggcc agctcgtttt tctctctgaa    2160
tctggctggt ttaggaggag cctgggttta tcgacgagat ctggagtatc tattcttttc    2220
cactgcttgc agtctccaat gtaggcagtg taaaggtata gtaaaatgat tttaggagtc    2280
agaaccaaat tgccaatatg ctccatggct cctaaaggaa aataaaatgg aagtttttaa    2340
aaaaaaaaaa aaaaaa                                                    2356
```

<210> 73

<211> 1646

<212> DNA

<213> Homo sapiens

<400> 73

```
gtggaatgtc atcagttaag gctattttca tttcttttgt ggatcttcag ttgcttcagg      60
ccatctggat gtatacatgc aggtcacagg gaatatgatg gcttagcttg ggttcagagg     120
cctgacacct caggctgcca aatgtggaag atttaaatac ttgaaccaat accctcctcc     180
caaaaactga aattggcttc tgtttctgag ttggtccagg cgcaatgttc agcgtatttg     240
aggaaatcac aagaattgta gttaaggaga tggatgctgg aggggatatg attgccgtta     300
gaagccttgt tgatgctgat agattccgct gcttccatct ggtgggggag aagagaactt     360
tctttggatg ccggcactac acaacaggcc tcaccctgat ggacattctg gacacacatg     420
gggacaagtg gttagatgaa ctggattctg ggctccaagg tcaaaaggct gagtttcaaa     480
ttctggataa tgtagactca acgggagagt tgatagtgag attacccaaa gaaataacaa     540
tttcaggcag tttccagggc ttccaccatc agaaaatcaa gatatcggag aaccggatat     600
cccagcagta tctggctacc cttgaaaaca ggaagctgaa gagggaacta ccctttttcat    660
tccgatcaat taatacgaga gaaaacctgt atctggtgac agaaactctg gagacggtaa     720
aggaggaaac cctgaaaagc gaccggcaat ataaattttg gagccagatc tctcagggcc     780
atctcagcta taaacacaag ggccaaaggg aagtgaccat ccccccaaat cgggtcctga     840
gctatcgagt aaagcagctt gtcttcccca acaaggagac gatgagaaag tctttgggtt     900
cggaggattc cagaaacatg aaggagaagt tggaggacat ggagagtgtc ctcaaggacc     960
tgacagagga gaagagaaaa gatgtgctaa actccctcgc taagtgcctc ggcaaggagg    1020
atattcggca ggatctagag caaagagtat ctgaggtcct gatttccggg gagctacaca    1080
tggaggaccc agacaagcct ctcctaagca gcctttttaa tgctgctggg gtcttggtag    1140
aagcgcgtgc aaaagccatt ctggacttcc tggatgccct gctagagctg tctgaagagc    1200
agcagtttgt ggctgaggcc ctggagaagg ggacccttcc tctgttgaag gaccaggtga    1260
aatctgtcat ggagcagaac tgggatgagc tggccagcag tcctcctgac atggactatg    1320
accctgaggc acgaattctc tgtgcgctgt atgttgttgt ctctatcctg ctggagctgg    1380
ctgaggggcc tacctctgtc tcttcctaac tacaaaagcc ctttctcccc acaagcctct    1440
gggttttccc tttaccagtc tgtcctcact gccatcgcca ctaccatcct gtcaccagtg    1500
ggacctcttt aaaacaagca gccaaccatt ctttgatgta tcccattcgc tccatgttaa    1560
catccaaaac cagcctggat ttcatacatg gacttctgat taaaagtggc aggttgtgca    1620
tgttaaaaaa aaaaaaaaaa aaaaaa                                         1646
```

<210> 74

<211> 3340

<212> DNA

<213> Homo sapiens

<400> 74

```
cgggcgccca gagacagcgc cgcctcagat atcctgctgg atgacattgt ccttacccat     60
tctctcttcc tcccgacgga gaaatttctg caggagctac accagtactt tgttcgggca    120
ggaggcatgg agggccctga agggctgggc cggaagcaag cctgtctagc catgcttctc    180
catttcttgg acacctacca ggggctgctt caagaggaag agggggccgg ccacatcatc    240
aaggatctat acctgctaat tatgaaggac gagtcccttt accagggcct ccgagaggac    300
actctgaggc tgcaccagct ggtggagacg gtggaactaa agattccaga ggagaaccag    360
ccacccagca agcaggtgaa gccactcttc cgccacttcc gccggataga ctcctgtctg    420
cagacccggg tggccttccg gggctctgat gagatcttct gccgtgtata catgcctgac    480
cactcttatg tgaccatacg cagccgcctt tcagcatctg tgcaggacat tctgggctct    540
gtgacggaga aacttcaata ttcagaggag cccgcggggc gtgaggattc cctcatcctg    600
gtagctgtgt cctcctctgg agagaaggtc cttctccagc ccactgagga ctgtgttttc    660
accgcactgg gcatcaacag ccacctgttt gcctgtactc gggacagcta tgaggctctg    720
gtgcccctcc ccgaggagat ccaggtctcc cctggagaca cagagatcca ccgagtggag    780
cctgaggacg ttgccaacca cctaactgcc ttccactggg agctgttccg atgtgtgcat    840
gagctggagt tcgtggacta cgtgttccac ggggagcgcg gccgccggga gacggccaac    900
ttggagctgc tgctgcagcg ctgcagcgag gtcacgcact gggtggccac cgaagtgctg    960
ctctgcgagg ccccgggcaa gcgcgcgcag ctgctcaaga agttcatcaa gatcgcggcc   1020
ctctgcaagc agaaccagga cctgctgtct ttctacgccg tggtcatggg gctggacaac   1080
gccgctgtca gccgccttcg actcacctgg gagaagctgc cagggaaatt caagaacttg   1140
tttcgcaaat ttgagaacct gacggacccc tgcaggaacc acaaaagcta ccgagaagtg   1200
atctccaaaa tgaagccccc tgtgattccc ttcgtgcctc tgatcctcaa agacctgact   1260
ttcctgcacg aagggagtaa gacccttgta gatggtttgg tgaacatcga gaagctgcat   1320
tcagtggccg aaaaagtgag gacaatccgc aaataccgga gccggcccct ttgcctggac   1380
atggaggcat cccccaatca cctgcagacc aaggcctatg tgcgccagtt tcaggtcatc   1440
gacaaccaga acctcctctt cgagctctcc tacaagctgg aggcaaacag tcagtgagag   1500
tggaggctcc agtcagaccc gccagatcct tgggcacctg gcactcaagc actttgcacg   1560
atgtctcaac caacatctga catctttccc gtggagcaac ttcctgctcc acgggaaaga   1620
ggtcgatgga tttacccctg gacccataag tctgttcatc ctgctgaagt cccctcccca   1680
ttgctccttc aagccaaaac tacactttgc tggttcctgt cccctctgag aaaggggata   1740
gaaagctcct tcctctatgt cctcccatcg agatctgttc tggggatgga gcttccaact   1800
tcctcttgca gcaggaaaga atgctgctca cccttctgtc ttgcagagtg ggattgtggg   1860
agggattggc agccttcttc tccaccacct gtccagcttc ttcctggtca gggctgggac   1920
ccccaggaat attatgttgc cgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg   1980
tgtgtcttct tttagggagc aggagtgcat ctggtaattg agggtggatg ttgtgtgtgc   2040
tggggagggg tccttctgtt tggtgctacc cttgtctact ctgcccctgg atggtgcggg   2100
gtgctttctc cacccccaca ctccctgctc agctcctcgt gctgccctgc atgcccaggc   2160
ttgtgagcca aggtgctttt tggggcaggg agtagcagca ggtgggaggg gttacccatc   2220
agcccttgca agtcccccac tcaggcctct ggaaggtcca gggatgggct ctgatgagag   2280
ggtaaaagat gctcagggaa acacaggcct cagctgccta gaggaccctc cccctgcctt   2340
gcagtgggct cgggtagagc agtatcagga gctagggttg tctgctgccc acactcctgc   2400
tttttgggat atctaactgc taaggaggga gttgacatcc cccttctggc tcatgtgtct   2460
gacaccaaca acatggtctc cgtccctctc tcttagactc tcccttttgtc ctccccatag   2520
agctggggtg gggtggatcc ctatactggg gcaggcagcc ccaaagtggg ggaggggat    2580
ggcagagact gtaaaggcgc cactggactc tggcaaggcc tttattacct ttactccctc   2640
cctctcccat caccagcctc aaggcctgag gggtgcaggg gctcctggca gctactgggt   2700
gaggtttcct ggcacagact cacccttctt tctggcacca ctctttccct tttgaagaga   2760
cagcaacagc cgtagcaaaa gcagctgctg ctcctgctat gagggtgtat atattttta    2820
cccaaagctc tggaattgta catttatttt ttaaaactca aagagggaaa gagccttgta   2880
tcatatgtga acattgtatc ataggtaatg ttgtacagac ccttttatac agtgatctgt   2940
cttgttcctg cagcaaaaat cctctatgga cataggaggt gctgtgtccc atgccttctt   3000
gccctgacag tgtcccatgg gccccccttct gctccctgcc ccctccctgc tactgctgat  3060
gcactgtcct ctccctgcag cccctggctt cccagccttc ctcctgaccc cttccaacag   3120
ccttggaact ccagctgcca ccaccctctg ggtcggacac tgggacccac tggcccagtc   3180
ttggctgctg cttaccccta gccttgatgc ctgcccaggg acccccagcc ccctcccgtt   3240
gccctgcagc tttaacagag tgaaccatgt gtattgtaca ggcgcggttg tcattgcaga   3300
aaccgctggg tggagaagaa gccgataaag tctatgaatc                         3340
```

<210> 75

<211> 4005

<212> DNA

<213> Homo sapiens

<400> 75
gggcaacagt ctgcccacct gtggacacca gatcctggga gctcctggtt agcaagtgag 60
atctctggga tgtcagtgag gctggttgaa gaccagaggt aaactgcaga ggtcaccacc 120
cccaccatgt cccaggtgat gtccagccca ctgctggcag gaggccatgc tgtcagcttg 180
gcgccttgtg atgagcccag gaggaccctg cacccagcac ccagccccag cctgccaccc 240
cagtgttctt actacaccac ggaaggctgg ggagcccagg ccctgatggc ccccgtgccc 300
tgcatggggc cccctggccg actccagcaa gccccacagg tggaggccaa agccacctgc 360
ttcctgccgt cccctggtga gaaggccttg gggaccccag aggaccttga ctcctacatt 420
gacttctcac tggagagcct caatcagatg atcctggaac tggaccccac cttccagctg 480
cttcccccag ggactggggg ctcccaggct gagctggccc agagcaccat gtcaatgaga 540
aagaaggagg aatctgaagc cttggacata aagtacatcg aggtgacctc cgccagatca 600
aggtgccacg attggcccca gcactgctcc agcccctctg tcaccccgcc cttcggctcc 660
cctcgcagtg gtggcctcct cctttccaga gacgtccccc gagagacacg aagcagcagt 720
gagagcctca tcttctctgg gaaccagggc agggggcacc agcgccctct gccccctca 780
gagggtctct cccctcgacc cccaaattcc cccagcatct caatcccttg catggggagc 840
aaggcctcga gcccccatgg tttgggctcc ccgctggtgg cttctccaag actggagaag 900
cggctgggag gcctggcccc acagcggggc agcaggatct ctgtgctgtc agccagccca 960
gtgtctgatg tcagctatat gtttggaagc agccagtccc tcctgcactc cagcaactcc 1020
agccatcagt catcttccag atccttggaa agtccagcca actcttcctc cagcctccac 1080
agccttggct cagtgtccct gtgtacaaga cccagtgact tccaggctcc cagaaacccc 1140
accctaacca tgggccaacc cagaacaccc cactctccac cactggccaa agaacatgcc 1200
agcatctgcc ccccatccat caccaactcc atggtggaca tacccattgt gctgatcaac 1260
ggctgcccag aaccagggtc ttctccaccc cagcggaccc caggacacca gaactccgtt 1320
caacctggag ctgcttctcc cagcaacccc tgtccagcca ccaggagcaa cagccagacc 1380
ctgtcagatg ccccctttac cacatgccca gagggtcccg ccagggacat gcagcccacc 1440
atgaagttcg tgatggacac atctaaatac tggtttaagc caaacatcac ccgagagcaa 1500
gcaatcgagc tgctgaggaa ggaggagcca ggggcttttg tcataaggga cagctcttca 1560
taccgaggct ccttcggcct ggccctgaag gtgcaggagg ttcccgcgtc tgctcagaat 1620
cgaccaggtg aggacagcaa tgacctcatc cgacacttcc tcatcgagtc gtctgccaaa 1680
ggagtgcatc tcaaaggagc agatgaggag ccctactttg ggagcctctc tgccttcgtg 1740
tgccagcatt ccatcatggc cctggccctg ccctgcaaac tcaccatccc acagagagaa 1800
ctgggaggtg cagatggggc ctcggactct acagacagcc cagcctcctg ccagaagaaa 1860
tctgcgggct gccacaccct gtacctgagc tcagtgagcg tggagaccct gactggagcc 1920
ctggccgtgc agaaagccat ctccaccacc tttgagaggg acatcctccc cacgcccacc 1980
gtggtccact tcgaagtcac agagcagggc atcactctga ctgatgtcca gaggaaggtg 2040
tttttccggc gccattaccc actcaccacc ctccgcttct gtggtatgga ccctgagcaa 2100
cggaagtggc agaagtactg caaaccctcc tggatctttg ggtttgtggc caagagccag 2160
acagagcctc aggagaacgt atgccacctc tttgcggagt atgacatggt ccagccagcc 2220
tcgcaggtca tcggcctggt gactgctctg ctgcaggacg cagaaaggat gtaggggaga 2280
gactgcctgt gcacctaacc aacacctcca ggggctcgct aaggagcccc cctccacccc 2340
ctgaatgggt gtggcttgtg gccatattga cagaccaatc tatgggacta gggggattgg 2400
catcaagttg acacccttga acctgctatg gccttcagca gtcaccatca tccagacccc 2460
ccgggcctca gtttcctcaa tcatagaaga agaccaatag acaagatcag ctgttcttag 2520
atgctggtgg gcatttgaac atgctcctcc atgattctga agcatgcaca cctctgaaga 2580
cccctgcatg aaaataacct ccaaggaccc tctgacccca tcgacctggg ccctgcccac 2640
acaacagtct gagcaagaga cctgcagccc ctgtttcgtg gcagacagca ggtgcctggc 2700
ggtgacccac ggggctcctg gcttgcagct ggtgatggtc aagaactgac tacaaaacag 2760
gaatggatag actctatttc cttccatatc tgttcctctg ttcctttttcc cactttctgg 2820
gtggcttttt gggtccaccc agccaggatg ctgcaggcca agctgggtgt ggtatttagg 2880
gcagctcagc aggggaact tgtccccatg gtcagaggag acccagctgt cctgcacccc 2940
cttgcagatg agtatcaccc catcttttct ttccacttgg tttttatttt tattttttt 3000
gagacagagt ctcactgtca cccaggctga actgcagtgg tgtgatctag gctcactgca 3060
acctccacct cccaggttca agcaattatc ctgcctcagg ctcccgagta gctgggatta 3120
caggcatgtg caactcaccc agctaatttt gtatttttag tagagacagg gtttcaccat 3180
gttggccagg ctggtcttga actcctgacc gcaggtaatc cacctgcttc ggcctcccaa 3240
agtgctggga ttacaggcgc aagccaccca gcccagcttc tttccattcc ttgataggcg 3300

agtattccaa agctggtatc gtagctgccc taatgttgca tattaggcgg cgggggcaga 3360
gataagggcc atctctctgt gattctgcct cagctcctgt cttgctgagc cctcccccaa 3420
cccacgctcc aacacacaca cacacacaca cacacacaca cacacacaca cacacacaca 3480
cacgcccctc tactgctatg tggcttcaac cagcctcaca gccacacggg ggaagcagag 3540
agtcaagaat gcaaagaggc cgcttcccta agaggcttgg aggagctggg ctctatccca 3600
cacccacccc cacccccaccc ccacccagcc tccagaagct ggaaccattt ctcccgcagg 3660
cctgagttcc taaggaaacc accctaccgg ggtggaaggg agggtcaggg aagaaaccca 3720
ctcttgctct acgaggagca agtgcctgcc ccctcccagc agccagccct gccaaagttg 3780
cattatcttt ggccaaggct gggcctgacg gttatgattt cagccctggg cctgcaggag 3840
aggctgagat cagcccaccc agccagtggt cgagcactgc cccgccgcca aagtctgcag 3900
aatgtgagat gaggttctca aggtcacagg ccccagtccc agcctggggg ctggcagagg 3960
cccccatata ctctgctaca gctcctatca tgaaaaataa aatgt 4005

<210> 76

<211> 1093

<212> PRT

<213> Homo sapiens

<400> 76

Met Lys Glu Met Val Gly Gly Cys Cys Val Cys Ser Asp Glu Arg Gly
1 5 10 15
Trp Ala Glu Asn Pro Leu Val Tyr Cys Asp Gly His Ala Cys Ser Val
20 25 30
Ala Val His Gln Ala Cys Tyr Gly Ile Val Gln Val Pro Thr Gly Pro
35 40 45
Trp Phe Cys Arg Lys Cys Glu Ser Gln Glu Arg Ala Ala Arg Val Arg
50 55 60
Cys Glu Leu Cys Pro His Lys Asp Gly Ala Leu Lys Arg Thr Asp Asn
65 70 75 80
Gly Gly Trp Ala His Val Val Cys Ala Leu Tyr Ile Pro Glu Val Gln
85 90 95
Phe Ala Asn Val Leu Thr Met Glu Pro Ile Val Leu Gln Tyr Val Pro
100 105 110
His Asp Arg Phe Asn Lys Thr Cys Tyr Ile Cys Glu Glu Thr Gly Arg
115 120 125
Glu Ser Lys Ala Ala Ser Gly Ala Cys Met Thr Cys Asn Arg His Gly
130 135 140
Cys Arg Gln Ala Phe His Val Thr Cys Ala Gln Met Ala Gly Leu Leu
145 150 155 160
Cys Glu Glu Glu Val Leu Glu Val Asp Asn Val Lys Tyr Cys Gly Tyr
165 170 175
Cys Lys Tyr His Phe Ser Lys Met Lys Thr Ser Arg His Ser Ser Gly
180 185 190
Gly Gly Gly Gly Gly Ala Gly Gly Gly Gly Gly Ser Met Gly Gly Gly
195 200 205
Gly Ser Gly Phe Ile Ser Gly Arg Arg Ser Arg Ser Ala Ser Pro Ser
210 215 220
Thr Gln Gln Glu Lys His Pro Thr His His Glu Arg Gly Gln Lys Lys
225 230 235 240
Ser Arg Lys Asp Lys Glu Arg Leu Lys Gln Lys His Lys Lys Arg Pro
245 250 255
Glu Ser Pro Pro Ser Ile Leu Thr Pro Pro Val Val Pro Thr Ala Asp
260 265 270
Lys Val Ser Ser Ser Ala Ser Ser Ser Ser His His Glu Ala Ser Thr
275 280 285
Gln Glu Thr Ser Glu Ser Ser Arg Glu Ser Lys Gly Lys Lys Ser Ser
290 295 300
Ser His Ser Leu Ser His Lys Gly Lys Lys Leu Ser Ser Gly Lys Gly
305 310 315 320
Val Ser Ser Phe Thr Ser Ala Ser Ser Ser Ser Ser Ser Ser Ser Ser
325 330 335

Ser Ser Gly Gly Pro Phe Gln Pro Ala Val Ser Ser Leu Gln Ser Ser
340 345 350

Pro Asp Phe Ser Ala Phe Pro Lys Leu Glu Gln Pro Glu Glu Asp Lys
355 360 365

Tyr Ser Lys Pro Thr Ala Pro Ala Pro Ser Ala Pro Pro Ser Pro Ser
370 375 380

Ala Pro Glu Pro Pro Lys Ala Asp Leu Phe Glu Gln Lys Val Val Phe
385 390 395 400

Ser Gly Phe Gly Pro Ile Met Arg Phe Ser Thr Thr Thr Ser Ser Ser
405 410 415

Gly Arg Ala Arg Ala Pro Ser Pro Gly Asp Tyr Lys Ser Pro His Val
420 425 430

Thr Gly Ser Gly Ala Ser Ala Gly Thr His Lys Arg Met Pro Ala Leu
435 440 445

Ser Ala Thr Pro Val Pro Ala Asp Glu Thr Pro Glu Thr Gly Leu Lys
450 455 460

Glu Lys Lys His Lys Ala Ser Lys Arg Ser Arg His Gly Pro Gly Arg
465 470 475 480

Pro Lys Gly Ser Arg Asn Lys Glu Gly Thr Gly Gly Pro Ala Ala Pro
485 490 495

Ser Leu Pro Ser Ala Gln Leu Ala Gly Phe Thr Ala Thr Ala Ala Ser
500 505 510

Pro Phe Ser Gly Gly Ser Leu Val Ser Ser Gly Leu Gly Gly Leu Ser
515 520 525

Ser Arg Thr Phe Gly Pro Ser Gly Ser Leu Pro Ser Leu Ser Leu Glu
530 535 540

Ser Pro Leu Leu Gly Ala Gly Ile Tyr Thr Ser Asn Lys Asp Pro Ile
545 550 555 560

Ser His Ser Gly Gly Met Leu Arg Ala Val Cys Ser Thr Pro Leu Ser
565 570 575

Ser Ser Leu Leu Gly Pro Pro Gly Thr Ser Ala Leu Pro Arg Leu Ser
580 585 590

Arg Ser Pro Phe Thr Ser Thr Leu Pro Ser Ser Ser Ala Ser Ile Ser
595 600 605

Thr Thr Gln Val Phe Ser Leu Ala Gly Ser Thr Phe Ser Leu Pro Ser
610 615 620

Thr His Ile Phe Gly Thr Pro Met Gly Ala Val Asn Pro Leu Leu Ser
625 630 635 640

Gln Ala Glu Ser Ser His Thr Glu Pro Asp Leu Glu Asp Cys Ser Phe
645 650 655

Arg Cys Arg Gly Thr Ser Pro Gln Glu Ser Leu Ser Ser Met Ser Pro
660 665 670

Ile Ser Ser Leu Pro Ala Leu Phe Asp Gln Thr Ala Ser Ala Pro Cys
675 680 685

Gly Gly Gly Gln Leu Asp Pro Ala Ala Pro Gly Thr Thr Asn Met Glu
690 695 700

Gln Leu Leu Glu Lys Gln Gly Asp Gly Glu Ala Gly Val Asn Ile Val
705 710 715 720

Glu Met Leu Lys Ala Leu His Ala Leu Gln Lys Glu Asn Gln Arg Leu
725 730 735

Gln Glu Gln Ile Leu Ser Leu Thr Ala Lys Lys Glu Arg Leu Gln Ile
740 745 750

Leu Asn Val Gln Leu Ser Val Pro Phe Pro Ala Leu Pro Ala Ala Leu
755 760 765

Pro Ala Ala Asn Gly Pro Val Pro Gly Pro Tyr Gly Leu Pro Pro Gln
770 775 780

Ala Gly Ser Ser Asp Ser Leu Ser Thr Ser Lys Ser Pro Pro Gly Lys
785 790 795 800

Ser Ser Leu Gly Leu Asp Asn Ser Leu Ser Thr Ser Ser Glu Asp Pro
805 810 815

His Ser Gly Cys Pro Ser Arg Ser Ser Ser Ser Leu Ser Phe His Ser
820 825 830

Thr Pro Pro Pro Leu Pro Leu Leu Gln Gln Ser Pro Ala Thr Leu Pro
835 840 845

Leu Ala Leu Pro Gly Ala Pro Ala Pro Leu Pro Pro Gln Pro Gln Asn
850 855 860

Gly Leu Gly Arg Ala Pro Gly Ala Ala Gly Leu Gly Ala Met Pro Met
865 870 875 880
Ala Glu Gly Leu Leu Gly Gly Leu Ala Gly Ser Gly Gly Leu Pro Leu
885 890 895
Asn Gly Leu Leu Gly Gly Leu Asn Gly Ala Ala Ala Pro Asn Pro Ala
900 905 910
Ser Leu Ser Gln Ala Gly Gly Ala Pro Thr Leu Gln Leu Pro Gly Cys
915 920 925
Leu Asn Ser Leu Thr Glu Gln Gln Arg His Leu Leu Gln Gln Gln Glu
930 935 940
Gln Gln Leu Gln Gln Leu Gln Gln Leu Leu Ala Ser Pro Gln Leu Thr
945 950 955 960
Pro Glu His Gln Thr Val Val Tyr Gln Met Ile Gln Gln Ile Gln Gln
965 970 975
Lys Arg Glu Leu Gln Arg Leu Gln Met Ala Gly Gly Ser Gln Leu Pro
980 985 990
Met Ala Ser Leu Leu Ala Gly Ser Ser Thr Pro Leu Leu Ser Ala Gly
995 1000 1005
Thr Pro Gly Leu Leu Pro Thr Ala Ser Ala Pro Pro Leu Leu Pro
1010 1015 1020
Ala Gly Ala Leu Val Ala Pro Ser Leu Gly Asn Asn Thr Ser Leu
1025 1030 1035
Met Ala Ala Ala Ala Ala Ala Ala Ala Val Ala Ala Ala Gly Gly
1040 1045 1050
Pro Pro Val Leu Thr Ala Gln Thr Asn Pro Phe Leu Ser Leu Ser
1055 1060 1065
Gly Ala Glu Gly Ser Gly Gly Gly Pro Lys Gly Gly Thr Ala Asp
1070 1075 1080
Lys Gly Ala Ser Ala Asn Gln Glu Lys Gly
1085 1090

<210> 77

<211> 344

<212> PRT

<213> Homo sapiens

<400> 77
Met His Arg Thr Thr Arg Ile Lys Ile Thr Glu Leu Asn Pro His Leu
1 5 10 15
Met Cys Ala Leu Cys Gly Gly Tyr Phe Ile Asp Ala Thr Thr Ile Val
20 25 30
Glu Cys Leu His Ser Phe Cys Lys Thr Cys Ile Val Arg Tyr Leu Glu
35 40 45
Thr Asn Lys Tyr Cys Pro Met Cys Asp Val Gln Val His Lys Thr Arg
50 55 60
Pro Leu Leu Ser Ile Arg Ser Asp Lys Thr Leu Gln Asp Ile Val Tyr
65 70 75 80
Lys Leu Val Pro Gly Leu Phe Lys Asp Glu Met Lys Arg Arg Arg Asp
85 90 95
Phe Tyr Ala Ala Tyr Pro Leu Thr Glu Val Pro Asn Gly Ser Asn Glu
100 105 110
Asp Arg Gly Glu Val Leu Glu Gln Glu Lys Gly Ala Leu Ser Asp Asp
115 120 125
Glu Ile Val Ser Leu Ser Ile Glu Phe Tyr Glu Gly Ala Arg Asp Arg
130 135 140
Asp Glu Lys Lys Gly Pro Leu Glu Asn Gly Asp Gly Asp Lys Glu Lys
145 150 155 160
Thr Gly Val Arg Phe Leu Arg Cys Pro Ala Ala Met Thr Val Met His
165 170 175
Leu Ala Lys Phe Leu Arg Asn Lys Met Asp Val Pro Ser Lys Tyr Lys
180 185 190

Val Glu Val Leu Tyr Glu Asp Glu Pro Leu Lys Glu Tyr Tyr Thr Leu
195 200 205
Met Asp Ile Ala Tyr Ile Tyr Pro Trp Arg Arg Asn Gly Pro Leu Pro
210 215 220
Leu Lys Tyr Arg Val Gln Pro Ala Cys Lys Arg Leu Thr Leu Ala Thr
225 230 235 240
Val Pro Thr Pro Ser Glu Gly Thr Asn Thr Ser Gly Ala Ser Glu Cys
245 250 255
Glu Ser Val Ser Asp Lys Ala Pro Ser Pro Ala Thr Leu Pro Ala Thr
260 265 270
Ser Ser Ser Leu Pro Ser Pro Ala Thr Pro Ser His Gly Ser Pro Ser
275 280 285
Ser His Gly Pro Pro Ala Thr His Pro Thr Ser Pro Thr Pro Pro Ser
290 295 300
Thr Ala Ser Gly Ala Thr Thr Ala Ala Asn Gly Gly Ser Leu Asn Cys
305 310 315 320
Leu Gln Thr Pro Ser Ser Thr Ser Arg Gly Arg Lys Met Thr Val Asn
325 330 335
Gly Ala Pro Val Pro Pro Leu Thr
340

<210> 78

<211> 416

<212> PRT

<213> Homo sapiens

<400> 78
Met Ser Ser Asn Cys Thr Ser Thr Thr Ala Val Ala Val Ala Pro Leu
1 5 10 15
Ser Ala Ser Lys Thr Lys Thr Lys Lys Lys His Phe Val Cys Gln Lys
20 25 30
Val Lys Leu Phe Arg Ala Ser Glu Pro Ile Leu Ser Val Leu Met Trp
35 40 45
Gly Val Asn His Thr Ile Asn Glu Leu Ser Asn Val Pro Val Pro Val
50 55 60
Met Leu Met Pro Asp Asp Phe Lys Ala Tyr Ser Lys Ile Lys Val Asp
65 70 75 80
Asn His Leu Phe Asn Lys Glu Asn Leu Pro Ser Arg Phe Lys Phe Lys
85 90 95
Glu Tyr Cys Pro Met Val Phe Arg Asn Leu Arg Glu Arg Phe Gly Ile
100 105 110
Asp Asp Gln Asp Tyr Gln Asn Ser Val Thr Arg Ser Ala Pro Ile Asn
115 120 125
Ser Asp Ser Gln Gly Arg Cys Gly Thr Arg Phe Leu Thr Thr Tyr Asp
130 135 140
Arg Arg Phe Val Ile Lys Thr Val Ser Ser Glu Asp Val Ala Glu Met
145 150 155 160
His Asn Ile Leu Lys Lys Tyr His Gln Phe Ile Val Glu Cys His Gly
165 170 175
Asn Thr Leu Leu Pro Gln Phe Leu Gly Met Tyr Arg Leu Thr Val Asp
180 185 190
Gly Val Glu Thr Tyr Met Val Val Thr Arg Asn Val Phe Ser His Arg
195 200 205
Leu Thr Val His Arg Lys Tyr Asp Leu Lys Gly Ser Thr Val Ala Arg
210 215 220
Glu Ala Ser Asp Lys Glu Lys Ala Lys Asp Leu Pro Thr Phe Lys Asp
225 230 235 240
Asn Asp Phe Leu Asn Glu Gly Gln Lys Leu His Val Gly Glu Glu Ser
245 250 255
Lys Lys Asn Phe Leu Glu Lys Leu Lys Arg Asp Val Glu Phe Leu Ala
260 265 270

Gln Leu Lys Ile Met Asp Tyr Ser Leu Leu Val Gly Ile His Asp Val
275 280 285
Asp Arg Ala Glu Gln Glu Glu Met Glu Val Glu Glu Arg Ala Glu Asp
290 295 300
Glu Glu Cys Glu Asn Asp Gly Val Gly Gly Asn Leu Leu Cys Ser Tyr
305 310 315 320
Gly Thr Pro Pro Asp Ser Pro Gly Asn Leu Leu Ser Phe Pro Arg Phe
325 330 335
Phe Gly Pro Gly Glu Phe Asp Pro Ser Val Asp Val Tyr Ala Met Lys
340 345 350
Ser His Glu Ser Ser Pro Lys Lys Glu Val Tyr Phe Met Ala Ile Ile
355 360 365
Asp Ile Leu Thr Pro Tyr Asp Thr Lys Lys Lys Ala Ala His Ala Ala
370 375 380
Lys Thr Val Lys His Gly Ala Gly Ala Glu Ile Ser Thr Val Asn Pro
385 390 395 400
Glu Gln Tyr Ser Lys Arg Phe Asn Glu Phe Met Ser Asn Ile Leu Thr
405 410 415

<210> 79

<211> 500

<212> PRT

<213> Homo sapiens

<400> 79
Met Arg Gly Glu Leu Trp Leu Leu Val Leu Val Leu Arg Glu Ala Ala
1 5 10 15
Arg Ala Leu Ser Pro Gln Pro Gly Ala Gly His Asp Glu Gly Pro Gly
20 25 30
Ser Gly Trp Ala Ala Lys Gly Thr Val Arg Gly Trp Asn Arg Arg Ala
35 40 45
Arg Glu Ser Pro Gly His Val Ser Glu Pro Asp Arg Thr Gln Leu Ser
50 55 60
Gln Asp Leu Gly Gly Gly Thr Leu Ala Met Asp Thr Leu Pro Asp Asn
65 70 75 80
Arg Thr Arg Val Val Glu Asp Asn His Ser Tyr Tyr Val Ser Arg Leu
85 90 95
Tyr Gly Pro Ser Glu Pro His Ser Arg Glu Leu Trp Val Asp Val Ala
100 105 110
Glu Ala Asn Arg Ser Gln Val Lys Ile His Thr Ile Leu Ser Asn Thr
115 120 125
His Arg Gln Ala Ser Arg Val Val Leu Ser Phe Asp Phe Pro Phe Tyr
130 135 140
Gly His Pro Leu Arg Gln Ile Thr Ile Ala Thr Gly Gly Phe Ile Phe
145 150 155 160
Met Gly Asp Val Ile His Arg Met Leu Thr Ala Thr Gln Tyr Val Ala
165 170 175
Pro Leu Met Ala Asn Phe Asn Pro Gly Tyr Ser Asp Asn Ser Thr Val
180 185 190
Val Tyr Phe Asp Asn Gly Thr Val Phe Val Val Gln Trp Asp His Val
195 200 205
Tyr Leu Gln Gly Trp Glu Asp Lys Gly Ser Phe Thr Phe Gln Ala Ala
210 215 220
Leu His His Asp Gly Arg Ile Val Phe Ala Tyr Lys Glu Ile Pro Met
225 230 235 240
Ser Val Pro Glu Ile Ser Ser Ser Gln His Pro Val Lys Thr Gly Leu
245 250 255
Ser Asp Ala Phe Met Ile Leu Asn Pro Ser Pro Asp Val Pro Glu Ser
260 265 270
Arg Arg Arg Ser Ile Phe Glu Tyr His Arg Ile Glu Leu Asp Pro Ser
275 280 285

Lys Val Thr Ser Met Ser Ala Val Glu Phe Thr Pro Leu Pro Thr Cys
290 295 300
Leu Gln His Arg Ser Cys Asp Ala Cys Met Ser Ser Asp Leu Thr Phe
305 310 315 320
Asn Cys Ser Trp Cys His Val Leu Gln Arg Cys Ser Ser Gly Phe Asp
325 330 335
Arg Tyr Arg Gln Glu Trp Met Asp Tyr Gly Cys Ala Gln Glu Ala Glu
340 345 350
Gly Arg Met Cys Glu Asp Phe Gln Asp Glu Asp His Asp Ser Ala Ser
355 360 365
Pro Asp Thr Ser Phe Ser Pro Tyr Asp Gly Asp Leu Thr Thr Thr Ser
370 375 380
Ser Ser Leu Phe Ile Asp Ser Leu Thr Thr Glu Asp Asp Thr Lys Leu
385 390 395 400
Asn Pro Tyr Ala Gly Gly Asp Gly Leu Gln Asn Asn Leu Ser Pro Lys
405 410 415
Thr Lys Gly Thr Pro Val His Leu Gly Thr Ile Val Gly Ile Val Leu
420 425 430
Ala Val Leu Leu Val Ala Ala Ile Ile Leu Ala Gly Ile Tyr Ile Asn
435 440 445
Gly His Pro Thr Ser Asn Ala Ala Leu Phe Phe Ile Glu Arg Arg Pro
450 455 460
His His Trp Pro Ala Met Lys Phe Arg Ser His Pro Asp His Ser Thr
465 470 475 480
Tyr Ala Glu Val Glu Pro Ser Gly His Glu Lys Glu Gly Phe Met Glu
485 490 495
Ala Glu Gln Cys
500

<210> 80

<211> 509

<212> PRT

<213> Homo sapiens

<400> 80
Met Glu Asp Ile Gln Thr Asn Ala Glu Leu Lys Ser Thr Gln Glu Gln
1 5 10 15
Ser Val Pro Ala Glu Ser Ala Ala Val Leu Asn Asp Tyr Ser Leu Thr
20 25 30
Lys Ser His Glu Met Glu Asn Val Asp Ser Gly Glu Gly Pro Ala Asn
35 40 45
Glu Asp Glu Asp Ile Gly Asp Asp Ser Met Lys Val Lys Asp Glu Tyr
50 55 60
Ser Glu Arg Asp Glu Asn Val Leu Lys Ser Glu Pro Met Gly Asn Ala
65 70 75 80
Glu Glu Pro Glu Ile Pro Tyr Ser Tyr Ser Arg Glu Tyr Asn Glu Tyr
85 90 95
Glu Asn Ile Lys Leu Glu Arg His Val Val Ser Phe Asp Ser Ser Arg
100 105 110
Pro Thr Ser Gly Lys Met Asn Cys Asp Val Cys Gly Leu Ser Cys Ile
115 120 125
Ser Phe Asn Val Leu Met Val His Lys Arg Ser His Thr Gly Glu Arg
130 135 140
Pro Phe Gln Cys Asn Gln Cys Gly Ala Ser Phe Thr Gln Lys Gly Asn
145 150 155 160
Leu Leu Arg His Ile Lys Leu His Thr Gly Glu Lys Pro Phe Lys Cys
165 170 175
His Leu Cys Asn Tyr Ala Cys Gln Arg Arg Asp Ala Leu Thr Gly His
180 185 190
Leu Arg Thr His Ser Val Glu Lys Pro Tyr Lys Cys Glu Phe Cys Gly
195 200 205

Arg Ser Tyr Lys Gln Arg Ser Ser Leu Glu Glu His Lys Glu Arg Cys
    210   215   220
Arg Thr Phe Leu Gln Ser Thr Asp Pro Gly Asp Thr Ala Ser Ala Glu
225   230   235   240
Ala Arg His Ile Lys Ala Glu Met Gly Ser Glu Arg Ala Leu Val Leu
                245   250   255
Asp Arg Leu Ala Ser Asn Val Ala Lys Arg Lys Ser Ser Met Pro Gln
            260   265   270
Lys Phe Ile Gly Glu Lys Arg His Cys Phe Asp Val Asn Tyr Asn Ser
        275   280   285
Ser Tyr Met Tyr Glu Lys Glu Ser Glu Leu Ile Gln Thr Arg Met Met
    290   295   300
Asp Gln Ala Ile Asn Asn Ala Ile Ser Tyr Leu Gly Ala Glu Ala Leu
305   310   315   320
Cys Pro Leu Val Gln Thr Pro Pro Ala Pro Thr Ser Glu Met Val Pro
                325   330   335
Val Ile Ser Ser Met Tyr Pro Ile Ala Leu Thr Arg Ala Glu Met Ser
            340   345   350
Asn Gly Ala Pro Gln Glu Leu Glu Arg Lys Ser Ile Leu Leu Pro Glu
        355   360   365
Lys Ser Val Pro Ser Glu Arg Gly Leu Ser Pro Asn Asn Ser Gly His
    370   375   380
Asp Ser Thr Asp Thr Asp Ser Asn His Glu Glu Arg Gln Asn His Ile
385   390   395   400
Tyr Gln Gln Asn His Met Val Leu Ser Arg Ala Arg Asn Gly Met Pro
                405   410   415
Leu Leu Lys Glu Val Pro Arg Ser Tyr Glu Leu Leu Lys Pro Pro Pro
            420   425   430
Ile Cys Pro Arg Asp Ser Val Lys Val Ile Asp Lys Glu Gly Glu Val
        435   440   445
Met Asp Val Tyr Arg Cys Asp His Cys Arg Val Leu Phe Leu Asp Tyr
    450   455   460
Val Met Phe Thr Ile His Met Gly Cys His Gly Phe Arg Asp Pro Phe
465   470   475   480
Glu Cys Asn Met Cys Gly Asp Arg Ser His Asp Arg Tyr Glu Phe Ser
                485   490   495
Ser His Ile Ala Arg Gly Glu His Arg Ser Leu Leu Lys
            500   505

<210> 81

<211> 440

<212> PRT

<213> Homo sapiens

<400> 81
Met Pro Ile Pro Pro Pro Pro Pro Pro Pro Pro Gly Pro Pro Pro Pro
1   5   10   15
Pro Thr Phe His Gln Ala Asn Thr Glu Gln Pro Lys Leu Ser Arg Asp
            20   25   30
Glu Gln Arg Gly Arg Gly Ala Leu Leu Gln Asp Ile Cys Lys Gly Thr
        35   40   45
Lys Leu Lys Lys Val Thr Asn Ile Asn Asp Arg Ser Ala Pro Ile Leu
    50   55   60
Glu Lys Pro Lys Gly Ser Ser Gly Gly Tyr Gly Ser Gly Gly Ala Ala
65   70   75   80
Leu Gln Pro Lys Gly Gly Leu Phe Gln Gly Gly Val Leu Lys Leu Arg
                85   90   95
Pro Val Gly Ala Lys Asp Gly Ser Glu Asn Leu Ala Gly Lys Pro Ala
            100   105   110
Leu Gln Ile Pro Ser Ser Arg Ala Ala Ala Pro Arg Pro Pro Val Ser
        115   120   125

Ala Ala Ser Gly Arg Pro Gln Asp Asp Thr Asp Ser Ser Arg Ala Ser
130 135 140
Leu Pro Glu Leu Pro Arg Met Gln Arg Pro Ser Leu Pro Asp Leu Ser
145 150 155 160
Arg Pro Asn Thr Thr Ser Ser Thr Gly Met Lys His Ser Ser Ser Ala
165 170 175
Pro Pro Pro Pro Pro Pro Gly Arg Arg Ala Asn Ala Pro Pro Thr Pro
180 185 190
Leu Pro Met His Ser Ser Lys Ala Pro Ala Tyr Asn Arg Glu Lys Pro
195 200 205
Leu Pro Pro Thr Pro Gly Gln Arg Leu His Pro Gly Arg Glu Gly Pro
210 215 220
Pro Ala Pro Pro Pro Val Lys Pro Pro Pro Ser Pro Val Asn Ile Arg
225 230 235 240
Thr Gly Pro Ser Gly Gln Ser Leu Ala Pro Pro Pro Pro Pro Tyr Arg
245 250 255
Gln Pro Pro Gly Val Pro Asn Gly Pro Ser Ser Pro Thr Asn Glu Ser
260 265 270
Ala Pro Glu Leu Pro Gln Arg His Asn Ser Leu His Arg Lys Thr Pro
275 280 285
Gly Pro Val Arg Gly Leu Ala Pro Pro Pro Pro Thr Ser Ala Ser Pro
290 295 300
Ser Leu Leu Ser Asn Arg Pro Pro Pro Pro Ala Arg Asp Pro Pro Ser
305 310 315 320
Arg Gly Ala Ala Pro Pro Pro Pro Pro Pro Val Ile Arg Asn Gly Ala
325 330 335
Arg Asp Ala Pro Pro Pro Pro Pro Pro Tyr Arg Met His Gly Ser Glu
340 345 350
Pro Pro Ser Arg Gly Lys Pro Pro Pro Pro Pro Ser Arg Thr Pro Ala
355 360 365
Gly Pro Pro Pro Pro Pro Pro Pro Pro Leu Arg Asn Gly His Arg Asp
370 375 380
Ser Ile Thr Thr Val Arg Ser Phe Leu Asp Asp Phe Glu Ser Lys Tyr
385 390 395 400
Ser Phe His Pro Val Glu Asp Phe Pro Ala Pro Glu Glu Tyr Lys His
405 410 415
Phe Gln Arg Ile Tyr Pro Ser Lys Thr Asn Arg Ala Ala Arg Gly Ala
420 425 430
Pro Pro Leu Pro Pro Ile Leu Arg
435 440

<210> 82

<211> 205

<212> PRT

<213> Homo sapiens

<400> 82
Met Ser Ile Met Ser Tyr Asn Gly Gly Ala Val Met Ala Met Lys Gly
1 5 10 15
Lys Asn Cys Val Ala Ile Ala Ala Asp Arg Arg Phe Gly Ile Gln Ala
20 25 30
Gln Met Val Thr Thr Asp Phe Gln Lys Ile Phe Pro Met Gly Asp Arg
35 40 45
Leu Tyr Ile Gly Leu Ala Gly Leu Ala Thr Asp Val Gln Thr Val Ala
50 55 60
Gln Arg Leu Lys Phe Arg Leu Asn Leu Tyr Glu Leu Lys Glu Gly Arg
65 70 75 80
Gln Ile Lys Pro Tyr Thr Leu Met Ser Met Val Ala Asn Leu Leu Tyr
85 90 95
Glu Lys Arg Phe Gly Pro Tyr Tyr Thr Glu Pro Val Ile Ala Gly Leu
100 105 110

Asp Pro Lys Thr Phe Lys Pro Phe Ile Cys Ser Leu Asp Leu Ile Gly
115 120 125
Cys Pro Met Val Thr Asp Asp Phe Val Val Ser Gly Thr Cys Ala Glu
130 135 140
Gln Met Tyr Gly Met Cys Glu Ser Leu Trp Glu Pro Asn Met Asp Pro
145 150 155 160
Asp His Leu Phe Glu Thr Ile Ser Gln Ala Met Leu Asn Ala Val Asp
165 170 175
Arg Asp Ala Val Ser Gly Met Gly Val Ile Val His Ile Ile Glu Lys
180 185 190
Asp Lys Ile Thr Thr Arg Thr Leu Lys Ala Arg Met Asp
195 200 205

<210> 83

<211> 190

<212> PRT

<213> Homo sapiens

<400> 83
Leu Thr Arg Ser Cys Ser Thr Cys Cys Pro Ala Val Ala Cys Leu Val
1 5 10 15
Gly Arg Gly Val Val Thr Ser Gly Ala Met His Gln Cys Trp Gly Glu
20 25 30
Glu Met Leu Gln Gly Met Leu Leu Trp Gly Trp Ala Thr Cys Pro Leu
35 40 45
Ser Asn Pro Gly Arg Trp Gly Arg Thr Val Gly Leu Gln His Pro Ala
50 55 60
Val Val Ser Ala Phe Arg Ala Leu Leu Leu Leu Met Leu Thr Val His
65 70 75 80
Val Ser Tyr Leu Ser Leu Ile Arg Phe Asp Tyr Gly Tyr Asn Leu Val
85 90 95
Ala Asn Val Ala Ile Gly Leu Val Asn Val Val Trp Trp Leu Ala Trp
100 105 110
Cys Leu Trp Asn Gln Arg Arg Leu Pro His Val Arg Lys Cys Val Val
115 120 125
Val Val Leu Leu Leu Gln Gly Leu Ser Leu Leu Glu Leu Leu Asp Phe
130 135 140
Pro Pro Leu Phe Trp Val Leu Asp Ala His Ala Ile Trp His Ile Ser
145 150 155 160
Thr Ile Pro Val His Val Leu Phe Phe Ser Phe Leu Glu Asp Asp Ser
165 170 175
Leu Tyr Leu Leu Lys Glu Ser Glu Asp Lys Phe Lys Leu Asp
180 185 190

<210> 84

<211> 368

<212> PRT

<213> Homo sapiens

<400> 84
Ala Pro Pro Pro Ala Ala Ser Gln Gly Glu Arg Met Ala Gly Leu Ala
1 5 10 15
Ala Arg Leu Val Leu Leu Ala Gly Ala Ala Ala Leu Ala Ser Gly Ser
20 25 30
Gln Gly Asp Arg Glu Pro Val Tyr Arg Asp Cys Val Leu Gln Cys Glu
35 40 45
Glu Gln Asn Cys Ser Gly Gly Ala Leu Asn His Phe Arg Ser Arg Gln
50 55 60

Pro Ile Tyr Met Ser Leu Ala Gly Trp Thr Cys Arg Asp Asp Cys Lys
65 70 75 80
Tyr Glu Cys Met Trp Val Thr Val Gly Leu Tyr Leu Gln Glu Gly His
85 90 95
Lys Val Pro Gln Phe His Gly Lys Trp Pro Phe Ser Arg Phe Leu Phe
100 105 110
Phe Gln Glu Pro Ala Ser Ala Val Ala Ser Phe Leu Asn Gly Leu Ala
115 120 125
Ser Leu Val Met Leu Cys Arg Tyr Arg Thr Phe Val Pro Ala Ser Ser
130 135 140
Pro Met Tyr His Thr Cys Val Ala Phe Ala Trp Val Ser Leu Asn Ala
145 150 155 160
Trp Phe Trp Ser Thr Val Phe His Thr Arg Asp Thr Asp Leu Thr Glu
165 170 175
Lys Met Asp Tyr Phe Cys Ala Ser Thr Val Ile Leu His Ser Ile Tyr
180 185 190
Leu Cys Cys Val Arg Thr Val Gly Leu Gln His Pro Ala Val Val Ser
195 200 205
Ala Phe Arg Ala Leu Leu Leu Leu Met Leu Thr Val His Val Ser Tyr
210 215 220
Leu Ser Leu Ile Arg Phe Asp Tyr Gly Tyr Asn Leu Val Ala Asn Val
225 230 235 240
Ala Ile Gly Leu Val Asn Val Val Trp Trp Leu Ala Trp Cys Leu Trp
245 250 255
Asn Gln Arg Arg Leu Pro His Val Arg Lys Cys Val Val Val Val Leu
260 265 270
Leu Leu Gln Gly Leu Ser Leu Leu Glu Leu Leu Asp Phe Pro Pro Leu
275 280 285
Phe Trp Val Leu Asp Ala His Ala Ile Trp His Ile Ser Thr Ile Pro
290 295 300
Val His Val Leu Phe Phe Ser Phe Leu Glu Asp Asp Ser Leu Tyr Leu
305 310 315 320
Leu Lys Glu Ser Glu Asp Lys Phe Lys Leu Val Glu Ala Asp Trp Ile
325 330 335
Phe Ala Leu Pro Leu Thr Pro Cys Pro Ser Leu Arg Glu Gly Ser Tyr
340 345 350
Ala Arg Thr Pro Thr Ser Gly Thr Arg Val Ala Cys Ala Ser Phe Phe
355 360 365

<210> 85

<211> 190

<212> PRT

<213> Homo sapiens

<400> 85

Leu Thr Arg Ser Cys Ser Thr Cys Cys Pro Ala Val Ala Cys Leu Val
1 5 10 15
Gly Arg Gly Val Val Thr Ser Gly Ala Met His Gln Cys Trp Gly Glu
20 25 30
Glu Met Leu Gln Gly Met Leu Leu Trp Gly Trp Ala Thr Cys Pro Leu
35 40 45
Ser Asn Pro Gly Arg Trp Gly Arg Thr Val Gly Leu Gln His Pro Ala
50 55 60
Val Val Ser Ala Phe Arg Ala Leu Leu Leu Leu Met Leu Thr Val His
65 70 75 80
Val Ser Tyr Leu Ser Leu Ile Arg Phe Asp Tyr Gly Tyr Asn Leu Val
85 90 95
Ala Asn Val Ala Ile Gly Leu Val Asn Val Val Trp Trp Leu Ala Trp
100 105 110
Cys Leu Trp Asn Gln Arg Arg Leu Pro His Val Arg Lys Cys Val Val
115 120 125

Val Val Leu Leu Leu Gln Gly Leu Ser Leu Leu Glu Leu Leu Asp Phe
130 135 140
Pro Pro Leu Phe Trp Val Leu Asp Ala His Ala Ile Trp His Ile Ser
145 150 155 160
Thr Ile Pro Val His Val Leu Phe Phe Ser Phe Leu Glu Asp Asp Ser
165 170 175
Leu Tyr Leu Leu Lys Glu Ser Glu Asp Lys Phe Lys Leu Asp
180 185 190

<210> 86

<211> 318

<212> PRT

<213> Homo sapiens

<400> 86
Met Ala Gly Leu Ala Ala Arg Leu Val Leu Leu Ala Gly Ala Ala Ala
1 5 10 15
Leu Ala Ser Gly Ser Gln Gly Asp Arg Glu Pro Val Tyr Arg Asp Cys
20 25 30
Val Leu Gln Cys Glu Glu Gln Asn Cys Ser Gly Gly Ala Leu Asn His
35 40 45
Phe Arg Ser Arg Gln Pro Ile Tyr Met Ser Leu Ala Gly Trp Thr Cys
50 55 60
Arg Asp Asp Cys Lys Tyr Glu Cys Met Trp Val Thr Val Gly Leu Tyr
65 70 75 80
Leu Gln Glu Gly His Lys Val Pro Gln Phe His Gly Lys Trp Pro Phe
85 90 95
Ser Arg Phe Leu Phe Phe Gln Glu Pro Ala Ser Ala Val Ala Ser Phe
100 105 110
Leu Asn Gly Leu Ala Ser Leu Val Met Leu Cys Arg Tyr Arg Thr Phe
115 120 125
Val Pro Ala Ser Ser Pro Met Tyr His Thr Cys Val Ala Phe Ala Trp
130 135 140
Val Ser Leu Asn Ala Trp Phe Trp Ser Thr Val Phe His Thr Arg Asp
145 150 155 160
Thr Asp Leu Gln Arg Lys Trp Thr Thr Ser Val Pro Pro Val Ser Tyr
165 170 175
Thr Gln Ser Thr Cys Ala Ala Ser Gly Pro Trp Gly Cys Ser Thr Gln
180 185 190
Leu Trp Ser Ser Ala Phe Arg Ala Leu Leu Leu Leu Met Leu Thr Val
195 200 205
His Val Ser Tyr Leu Ser Leu Ile Arg Phe Asp Tyr Gly Tyr Asn Leu
210 215 220
Val Ala Asn Val Ala Ile Gly Leu Val Asn Val Val Trp Trp Leu Ala
225 230 235 240
Trp Cys Leu Trp Asn Gln Arg Arg Leu Pro His Val Arg Lys Cys Val
245 250 255
Val Val Val Leu Leu Leu Gln Gly Leu Ser Leu Leu Glu Leu Leu Asp
260 265 270
Phe Pro Pro Leu Phe Trp Val Leu Asp Ala His Ala Ile Trp His Ile
275 280 285
Ser Thr Ile Pro Val His Val Leu Phe Phe Ser Phe Leu Glu Asp Asp
290 295 300
Ser Leu Tyr Leu Leu Lys Glu Ser Glu Asp Lys Phe Lys Leu
305 310 315

<210> 87

<211> 226

<212> PRT

<213> Homo sapiens

<400> 87
Met Ala Gly Leu Ala Ala Arg Leu Val Leu Leu Ala Gly Ala Ala Ala
1 5 10 15
Leu Ala Ser Gly Ser Gln Gly Asp Arg Glu Pro Val Tyr Arg Asp Cys
20 25 30
Val Leu Gln Cys Glu Glu Gln Asn Cys Ser Gly Gly Ala Leu Asn His
35 40 45
Phe Arg Ser Arg Gln Pro Ile Tyr Met Ser Leu Ala Gly Trp Thr Cys
50 55 60
Arg Asp Asp Cys Lys Tyr Glu Cys Met Trp Val Thr Val Gly Leu Tyr
65 70 75 80
Leu Gln Glu Gly His Lys Val Pro Gln Phe His Gly Lys Trp Pro Phe
85 90 95
Ser Arg Phe Leu Phe Phe Gln Glu Pro Ala Ser Ala Val Ala Ser Phe
100 105 110
Leu Asn Gly Leu Ala Ser Leu Val Met Leu Cys Arg Tyr Arg Thr Phe
115 120 125
Val Pro Ala Ser Ser Pro Met Tyr His Thr Cys Val Ala Phe Ala Trp
130 135 140
Val Ser Leu Asn Ala Trp Phe Trp Ser Thr Val Phe His Thr Arg Asp
145 150 155 160
Thr Asp Leu Thr Glu Lys Met Asp Tyr Phe Cys Ala Ser Thr Val Ile
165 170 175
Leu His Ser Ile Tyr Leu Cys Cys Val Arg Pro Gly Gln Arg Gly Val
180 185 190
Val Ala Gly Leu Val Pro Val Glu Pro Ala Ala Ala Ala Ser Arg Ala
195 200 205
Gln Val Arg Gly Gly Gly Leu Ala Ala Ala Gly Ala Val Pro Ala Arg
210 215 220
Ala Ala
225
<210> 88

<211> 320

<212> PRT

<213> Homo sapiens

<400> 88
Met Ala Gly Leu Ala Ala Arg Leu Val Leu Leu Ala Gly Ala Ala Ala
1 5 10 15
Leu Ala Ser Gly Ser Gln Gly Asp Arg Glu Pro Val Tyr Arg Asp Cys
20 25 30
Val Leu Gln Cys Glu Glu Gln Asn Cys Ser Gly Gly Ala Leu Asn His
35 40 45
Phe Arg Ser Arg Gln Pro Ile Tyr Met Ser Leu Ala Gly Trp Thr Cys
50 55 60
Arg Asp Asp Cys Lys Tyr Glu Cys Met Trp Val Thr Val Gly Leu Tyr
65 70 75 80
Leu Gln Glu Gly His Lys Val Pro Gln Phe His Gly Lys Trp Pro Phe
85 90 95
Ser Arg Phe Leu Phe Phe Gln Glu Pro Ala Ser Ala Val Ala Ser Phe
100 105 110
Leu Asn Gly Leu Ala Ser Leu Val Met Leu Cys Arg Tyr Arg Thr Phe
115 120 125
Val Pro Ala Ser Ser Pro Met Tyr His Thr Cys Val Ala Phe Ala Trp
130 135 140
Val Ser Leu Asn Ala Trp Phe Trp Ser Thr Val Phe His Thr Arg Asp
145 150 155 160

Thr Asp Leu Thr Glu Lys Met Asp Tyr Phe Cys Ala Ser Thr Val Ile
165 170 175
Leu His Ser Ile Tyr Leu Cys Cys Val Arg Thr Val Gly Leu Gln His
180 185 190
Pro Ala Val Val Ser Ala Phe Arg Ala Leu Leu Leu Leu Met Leu Thr
195 200 205
Val His Val Ser Tyr Leu Ser Leu Ile Arg Phe Asp Tyr Gly Tyr Asn
210 215 220
Leu Val Ala Asn Val Ala Ile Gly Leu Val Asn Val Val Trp Trp Leu
225 230 235 240
Ala Trp Cys Leu Trp Asn Gln Arg Arg Leu Pro His Val Arg Lys Cys
245 250 255
Val Val Val Val Leu Leu Leu Gln Gly Leu Ser Leu Leu Glu Leu Leu
260 265 270
Asp Phe Pro Pro Leu Phe Trp Val Leu Asp Ala His Ala Ile Trp His
275 280 285
Ile Ser Thr Ile Pro Val His Val Leu Phe Phe Ser Phe Leu Glu Asp
290 295 300
Asp Ser Leu Tyr Leu Leu Lys Glu Ser Glu Asp Lys Phe Lys Leu Asp
305 310 315 320

<210> 89

<211> 217

<212> PRT

<213> Homo sapiens

<400> 89
Ala Pro Pro Pro Ala Ala Ser Gln Gly Glu Arg Met Ala Gly Leu Ala
1 5 10 15
Ala Arg Leu Val Leu Leu Ala Gly Ala Ala Ala Leu Ala Ser Gly Ser
20 25 30
Gln Gly Asp Arg Glu Pro Val Tyr Arg Asp Cys Val Leu Gln Cys Glu
35 40 45
Glu Gln Asn Cys Ser Gly Gly Ala Leu Asn His Phe Arg Ser Arg Gln
50 55 60
Pro Ile Tyr Met Ser Leu Ala Gly Trp Thr Cys Arg Asp Asp Cys Lys
65 70 75 80
Tyr Glu Cys Met Trp Val Thr Val Gly Leu Tyr Leu Gln Glu Gly His
85 90 95
Lys Val Pro Gln Phe His Gly Lys Trp Pro Phe Ser Arg Phe Leu Phe
100 105 110
Phe Gln Glu Pro Ala Ser Ala Val Ala Ser Phe Leu Asn Gly Leu Ala
115 120 125
Ser Leu Val Met Leu Cys Arg Tyr Arg Thr Phe Val Pro Ala Ser Ser
130 135 140
Pro Met Tyr His Thr Cys Val Ala Phe Ala Trp Val Ser Leu Asn Ala
145 150 155 160
Trp Phe Trp Ser Thr Val Phe His Thr Arg Asp Thr Asp Leu Thr Glu
165 170 175
Lys Met Asp Tyr Phe Cys Ala Ser Thr Val Ile Leu His Ser Ile Tyr
180 185 190
Leu Cys Cys Val Ser Phe Leu Glu Asp Asp Ser Leu Tyr Leu Leu Lys
195 200 205
Glu Ser Glu Asp Lys Phe Lys Leu Asp
210 215

<210> 90

<211> 153

<212> PRT

<213> Homo sapiens

<400> 90

Met Asn Val Gly Thr Ala His Ser Glu Val Asn Pro Asn Thr Arg Val
1 5 10 15
Met Asn Ser Arg Gly Ile Trp Leu Ser Tyr Val Leu Ala Ile Gly Leu
20 25 30
Leu His Ile Val Leu Leu Ser Ile Pro Phe Val Ser Val Pro Val Val
35 40 45
Trp Thr Leu Thr Asn Leu Ile His Asn Met Gly Met Tyr Ile Phe Leu
50 55 60
His Thr Val Lys Gly Thr Pro Phe Glu Thr Pro Asp Gln Gly Lys Ala
65 70 75 80
Arg Leu Leu Thr His Trp Glu Gln Met Asp Tyr Gly Val Gln Phe Thr
85 90 95
Ala Ser Arg Lys Phe Leu Thr Ile Thr Pro Ile Val Leu Tyr Phe Leu
100 105 110
Thr Ser Phe Tyr Thr Lys Tyr Asp Gln Ile His Phe Val Leu Asn Thr
115 120 125
Val Ser Leu Met Ser Val Leu Ile Pro Lys Leu Pro Gln Leu His Gly
130 135 140
Val Arg Ile Phe Gly Ile Asn Lys Tyr
145 150

<210> 91

<211> 436

<212> PRT

<213> Homo sapiens

<400> 91

Met Arg Arg Asp Val Asn Gly Val Thr Lys Ser Arg Phe Glu Met Phe
1 5 10 15
Ser Asn Ser Asp Glu Ala Val Ile Asn Lys Lys Leu Pro Lys Glu Leu
20 25 30
Leu Leu Arg Ile Phe Ser Phe Leu Asp Val Val Thr Leu Cys Arg Cys
35 40 45
Ala Gln Val Ser Arg Ala Trp Asn Val Leu Ala Leu Asp Gly Ser Asn
50 55 60
Trp Gln Arg Ile Asp Leu Phe Asp Phe Gln Arg Asp Ile Glu Gly Arg
65 70 75 80
Val Val Glu Asn Ile Ser Lys Arg Cys Gly Gly Phe Leu Arg Lys Leu
85 90 95
Ser Leu Arg Gly Cys Leu Gly Val Gly Asp Asn Ala Leu Arg Thr Phe
100 105 110
Ala Gln Asn Cys Arg Asn Ile Glu Val Leu Asn Leu Asn Gly Cys Thr
115 120 125
Lys Thr Thr Asp Ala Thr Cys Thr Ser Leu Ser Lys Phe Cys Ser Lys
130 135 140
Leu Arg His Leu Asp Leu Ala Ser Cys Thr Ser Ile Thr Asn Met Ser
145 150 155 160
Leu Lys Ala Leu Ser Glu Gly Cys Pro Leu Leu Glu Gln Leu Asn Ile
165 170 175
Ser Trp Cys Asp Gln Val Thr Lys Asp Gly Ile Gln Ala Leu Val Arg
180 185 190
Gly Cys Gly Gly Leu Lys Ala Leu Phe Leu Lys Gly Cys Thr Gln Leu
195 200 205
Glu Asp Glu Ala Leu Lys Tyr Ile Gly Ala His Cys Pro Glu Leu Val
210 215 220
Thr Leu Asn Leu Gln Thr Cys Leu Gln Ile Thr Asp Glu Gly Leu Ile
225 230 235 240

Thr Ile Cys Arg Gly Cys His Lys Leu Gln Ser Leu Cys Ala Ser Gly
245 250 255
Cys Ser Asn Ile Thr Asp Ala Ile Leu Asn Ala Leu Gly Gln Asn Cys
260 265 270
Pro Arg Leu Arg Ile Leu Glu Val Ala Arg Cys Ser Gln Leu Thr Asp
275 280 285
Val Gly Phe Thr Thr Leu Ala Arg Asn Cys His Glu Leu Glu Lys Met
290 295 300
Asp Leu Glu Glu Cys Val Gln Ile Thr Asp Ser Thr Leu Ile Gln Leu
305 310 315 320
Ser Ile His Cys Pro Arg Leu Gln Val Leu Ser Leu Ser His Cys Glu
325 330 335
Leu Ile Thr Asp Asp Gly Ile Arg His Leu Gly Asn Gly Ala Cys Ala
340 345 350
His Asp Gln Leu Glu Val Ile Glu Leu Asp Asn Cys Pro Leu Ile Thr
355 360 365
Asp Ala Ser Leu Glu His Leu Lys Ser Cys His Ser Leu Glu Arg Ile
370 375 380
Glu Leu Tyr Asp Cys Gln Gln Ile Thr Arg Ala Gly Ile Lys Arg Leu
385 390 395 400
Arg Thr His Leu Pro Asn Ile Lys Val His Ala Tyr Phe Ala Pro Val
405 410 415
Thr Pro Pro Pro Ser Val Gly Gly Ser Arg Gln Arg Phe Cys Arg Cys
420 425 430
Cys Ile Ile Leu
435

<210> 92

<211> 204

<212> PRT

<213> Homo sapiens

<400> 92
Met Asp Pro Lys Asp Arg Lys Lys Ile Gln Phe Ser Val Pro Ala Pro
1 5 10 15
Pro Ser Gln Leu Asp Pro Arg Gln Val Glu Met Ile Arg Arg Arg Arg
20 25 30
Pro Thr Pro Ala Met Leu Phe Arg Leu Ser Glu His Ser Ser Pro Glu
35 40 45
Glu Glu Ala Ser Pro His Gln Arg Ala Ser Gly Glu Gly His His Leu
50 55 60
Lys Ser Lys Arg Pro Asn Pro Cys Ala Tyr Thr Pro Pro Ser Leu Lys
65 70 75 80
Ala Val Gln Arg Ile Ala Glu Ser His Leu Gln Ser Ile Ser Asn Leu
85 90 95
Asn Glu Asn Gln Ala Ser Glu Glu Glu Asp Glu Leu Gly Glu Leu Arg
100 105 110
Glu Leu Gly Tyr Pro Arg Glu Glu Asp Glu Glu Glu Glu Glu Asp Asp
115 120 125
Glu Glu Glu Glu Glu Glu Glu Asp Ser Gln Ala Glu Val Leu Lys Val
130 135 140
Ile Arg Gln Ser Ala Gly Gln Lys Thr Thr Cys Gly Gln Gly Leu Glu
145 150 155 160
Gly Pro Trp Glu Arg Pro Pro Pro Leu Asp Glu Ser Glu Arg Asp Gly
165 170 175
Gly Ser Glu Asp Gln Val Glu Asp Pro Ala Leu Ser Glu Pro Gly Glu
180 185 190
Glu Pro Gln Arg Pro Ser Pro Ser Glu Pro Gly Thr
195 200

<210> 93

<211> 115

<212> PRT

<213> Homo sapiens

<400> 93

Met Ser Gly Glu Pro Gly Gln Thr Ser Val Ala Pro Pro Pro Glu Glu
1 5 10 15
Val Glu Pro Gly Ser Gly Val Arg Ile Val Val Glu Tyr Cys Glu Pro
20 25 30
Cys Gly Phe Glu Ala Thr Tyr Leu Glu Leu Ala Ser Ala Val Lys Glu
35 40 45
Gln Tyr Pro Gly Ile Glu Ile Glu Ser Arg Leu Gly Gly Thr Gly Ala
50 55 60
Phe Glu Ile Glu Ile Asn Gly Gln Leu Val Phe Ser Lys Leu Glu Asn
65 70 75 80
Gly Gly Phe Pro Tyr Glu Lys Asp Leu Ile Glu Ala Ile Arg Arg Ala
85 90 95
Ser Asn Gly Glu Thr Leu Glu Lys Ile Thr Asn Ser Arg Pro Pro Cys
100 105 110
Val Ile Leu
115

<210> 94

<211> 144

<212> PRT

<213> Homo sapiens

<400> 94

Met Gly Ala Val Val Leu Cys Arg Pro Ser Pro Leu Asn Phe Leu Ile
1 5 10 15
Gln Thr Gly Thr Gly Gln Gly Leu Ser Cys Gly Ser His Met Trp Arg
20 25 30
Cys Glu Ala Thr Pro Cys Gly Val Cys Gly Glu Ser Pro Val Gly Ser
35 40 45
Leu Leu Lys Gln His Arg Gly Arg Gly Lys Thr Trp Pro Val Gly Thr
50 55 60
Val Ser Ala Cys Arg Glu Glu Ser Glu Ala Gly Ser Leu Ser Leu Gly
65 70 75 80
Trp Ser Leu Leu Pro Ser Pro Val Gly Leu Gly Ala Val Leu Ile Leu
85 90 95
Lys Arg Cys Gly Ser Leu Cys Pro Leu Pro Gly Val Gln Gly Asn Arg
100 105 110
Arg Gly His Trp Ala Cys Phe Leu Pro Pro Asp Pro Ala Ser Pro Thr
115 120 125
Pro Cys Ile Ile Gly Asn Phe His Leu Lys Ile Phe Leu Ser Lys Val
130 135 140

<210> 95

<211> 425

<212> PRT

<213> Homo sapiens

<400> 95

Met Gly Gly Gly Asp Leu Asn Leu Lys Lys Ser Trp His Pro Gln Thr
1 5 10 15

Leu Arg Asn Val Glu Lys Val Trp Lys Ala Glu Gln Lys His Glu Ala
20 25 30
Glu Arg Lys Lys Ile Glu Glu Leu Gln Arg Glu Leu Arg Glu Glu Arg
35 40 45
Ala Arg Glu Glu Met Gln Arg Tyr Ala Glu Asp Val Gly Ala Val Lys
50 55 60
Lys Lys Glu Glu Lys Leu Asp Trp Met Tyr Gln Gly Pro Gly Gly Met
65 70 75 80
Val Asn Arg Asp Glu Tyr Leu Leu Gly Arg Pro Ile Asp Lys Tyr Val
85 90 95
Phe Glu Lys Met Glu Glu Lys Glu Ala Gly Cys Ser Ser Glu Thr Gly
100 105 110
Leu Leu Pro Gly Ser Ile Phe Ala Pro Ser Gly Ala Asn Ser Leu Leu
115 120 125
Asp Met Ala Ser Lys Ile Arg Glu Asp Pro Leu Phe Ile Ile Arg Lys
130 135 140
Lys Glu Glu Glu Lys Lys Arg Glu Val Leu Asn Asn Pro Val Lys Met
145 150 155 160
Lys Lys Ile Lys Glu Leu Leu Gln Met Ser Leu Glu Lys Lys Glu Lys
165 170 175
Lys Lys Lys Lys Glu Lys Lys Lys Lys His Lys Lys His Lys His Arg
180 185 190
Ser Ser Ser Ser Asp Arg Ser Ser Ser Glu Asp Glu His Ser Ala Gly
195 200 205
Arg Ser Gln Lys Lys Met Ala Asn Ser Ser Pro Val Leu Ser Lys Val
210 215 220
Pro Gly Tyr Gly Leu Gln Val Arg Asn Ser Asp Arg Asn Gln Gly Leu
225 230 235 240
Gln Gly Pro Leu Thr Ala Glu Gln Lys Arg Gly His Gly Met Lys Asn
245 250 255
His Ser Arg Ser Arg Ser Ser Ser His Ser Pro Pro Arg His Ala Ser
260 265 270
Lys Lys Ser Thr Arg Glu Ala Gly Ser Arg Asp Arg Arg Ser Arg Ser
275 280 285
Leu Gly Arg Arg Ser Arg Ser Pro Arg Pro Ser Lys Leu His Asn Ser
290 295 300
Lys Val Asn Arg Arg Glu Thr Gly Gln Thr Arg Ser Pro Ser Pro Lys
305 310 315 320
Lys Glu Val Tyr Gln Arg Arg His Ala Pro Gly Tyr Thr Arg Lys Leu
325 330 335
Ser Ala Glu Glu Leu Glu Arg Lys Arg Gln Glu Met Met Glu Asn Ala
340 345 350
Lys Trp Arg Glu Glu Glu Arg Leu Asn Ile Leu Lys Arg His Ala Lys
355 360 365
Asp Glu Glu Arg Glu Gln Arg Leu Glu Lys Leu Asp Ser Arg Asp Gly
370 375 380
Lys Phe Ile His Arg Met Lys Leu Glu Ser Ala Ser Thr Ser Ser Leu
385 390 395 400
Glu Asp Arg Val Lys Arg Asn Ile Tyr Ser Leu Gln Arg Thr Ser Val
405 410 415
Ala Leu Glu Lys Asn Phe Met Lys Arg
420 425

<210> 96

<211> 394

<212> PRT

<213> Homo sapiens

<400> 96
Met Phe Ser Val Phe Glu Glu Ile Thr Arg Ile Val Val Lys Glu Met
1 5 10 15

Asp Ala Gly Gly Asp Met Ile Ala Val Arg Ser Leu Val Asp Ala Asp
20 25 30
Arg Phe Arg Cys Phe His Leu Val Gly Glu Lys Arg Thr Phe Phe Gly
35 40 45
Cys Arg His Tyr Thr Thr Gly Leu Thr Leu Met Asp Ile Leu Asp Thr
50 55 60
His Gly Asp Lys Trp Leu Asp Glu Leu Asp Ser Gly Leu Gln Gly Gln
65 70 75 80
Lys Ala Glu Phe Gln Ile Leu Asp Asn Val Asp Ser Thr Gly Glu Leu
85 90 95
Ile Val Arg Leu Pro Lys Glu Ile Thr Ile Ser Gly Ser Phe Gln Gly
100 105 110
Phe His His Gln Lys Ile Lys Ile Ser Glu Asn Arg Ile Ser Gln Gln
115 120 125
Tyr Leu Ala Thr Leu Glu Asn Arg Lys Leu Lys Arg Glu Leu Pro Phe
130 135 140
Ser Phe Arg Ser Ile Asn Thr Arg Glu Asn Leu Tyr Leu Val Thr Glu
145 150 155 160
Thr Leu Glu Thr Val Lys Glu Glu Thr Leu Lys Ser Asp Arg Gln Tyr
165 170 175
Lys Phe Trp Ser Gln Ile Ser Gln Gly His Leu Ser Tyr Lys His Lys
180 185 190
Gly Gln Arg Glu Val Thr Ile Pro Pro Asn Arg Val Leu Ser Tyr Arg
195 200 205
Val Lys Gln Leu Val Phe Pro Asn Lys Glu Thr Met Arg Lys Ser Leu
210 215 220
Gly Ser Glu Asp Ser Arg Asn Met Lys Glu Lys Leu Glu Asp Met Glu
225 230 235 240
Ser Val Leu Lys Asp Leu Thr Glu Glu Lys Arg Lys Asp Val Leu Asn
245 250 255
Ser Leu Ala Lys Cys Leu Gly Lys Glu Asp Ile Arg Gln Asp Leu Glu
260 265 270
Gln Arg Val Ser Glu Val Leu Ile Ser Gly Glu Leu His Met Glu Asp
275 280 285
Pro Asp Lys Pro Leu Leu Ser Ser Leu Phe Asn Ala Ala Gly Val Leu
290 295 300
Val Glu Ala Arg Ala Lys Ala Ile Leu Asp Phe Leu Asp Ala Leu Leu
305 310 315 320
Glu Leu Ser Glu Glu Gln Gln Phe Val Ala Glu Ala Leu Glu Lys Gly
325 330 335
Thr Leu Pro Leu Leu Lys Asp Gln Val Lys Ser Val Met Glu Gln Asn
340 345 350
Trp Asp Glu Leu Ala Ser Ser Pro Pro Asp Met Asp Tyr Asp Pro Glu
355 360 365
Ala Arg Ile Leu Cys Ala Leu Tyr Val Val Val Ser Ile Leu Leu Glu
370 375 380
Leu Ala Glu Gly Pro Thr Ser Val Ser Ser
385 390

<210> 97

<211> 456

<212> PRT

<213> Homo sapiens

<400> 97
Met Glu Gly Pro Glu Gly Leu Gly Arg Lys Gln Ala Cys Leu Ala Met
1 5 10 15
Leu Leu His Phe Leu Asp Thr Tyr Gln Gly Leu Leu Gln Glu Glu Glu
20 25 30
Gly Ala Gly His Ile Ile Lys Asp Leu Tyr Leu Leu Ile Met Lys Asp
35 40 45

Glu Ser Leu Tyr Gln Gly Leu Arg Glu Asp Thr Leu Arg Leu His Gln
50 55 60
Leu Val Glu Thr Val Glu Leu Lys Ile Pro Glu Glu Asn Gln Pro Pro
65 70 75 80
Ser Lys Gln Val Lys Pro Leu Phe Arg His Phe Arg Arg Ile Asp Ser
85 90 95
Cys Leu Gln Thr Arg Val Ala Phe Arg Gly Ser Asp Glu Ile Phe Cys
100 105 110
Arg Val Tyr Met Pro Asp His Ser Tyr Val Thr Ile Arg Ser Arg Leu
115 120 125
Ser Ala Ser Val Gln Asp Ile Leu Gly Ser Val Thr Glu Lys Leu Gln
130 135 140
Tyr Ser Glu Glu Pro Ala Gly Arg Glu Asp Ser Leu Ile Leu Val Ala
145 150 155 160
Val Ser Ser Ser Gly Glu Lys Val Leu Leu Gln Pro Thr Glu Asp Cys
165 170 175
Val Phe Thr Ala Leu Gly Ile Asn Ser His Leu Phe Ala Cys Thr Arg
180 185 190
Asp Ser Tyr Glu Ala Leu Val Pro Leu Pro Glu Glu Ile Gln Val Ser
195 200 205
Pro Gly Asp Thr Glu Ile His Arg Val Glu Pro Glu Asp Val Ala Asn
210 215 220
His Leu Thr Ala Phe His Trp Glu Leu Phe Arg Cys Val His Glu Leu
225 230 235 240
Glu Phe Val Asp Tyr Val Phe His Gly Glu Arg Gly Arg Arg Glu Thr
245 250 255
Ala Asn Leu Glu Leu Leu Leu Gln Arg Cys Ser Glu Val Thr His Trp
260 265 270
Val Ala Thr Glu Val Leu Leu Cys Glu Ala Pro Gly Lys Arg Ala Gln
275 280 285
Leu Leu Lys Lys Phe Ile Lys Ile Ala Ala Leu Cys Lys Gln Asn Gln
290 295 300
Asp Leu Leu Ser Phe Tyr Ala Val Val Met Gly Leu Asp Asn Ala Ala
305 310 315 320
Val Ser Arg Leu Arg Leu Thr Trp Glu Lys Leu Pro Gly Lys Phe Lys
325 330 335
Asn Leu Phe Arg Lys Phe Glu Asn Leu Thr Asp Pro Cys Arg Asn His
340 345 350
Lys Ser Tyr Arg Glu Val Ile Ser Lys Met Lys Pro Pro Val Ile Pro
355 360 365
Phe Val Pro Leu Ile Leu Lys Asp Leu Thr Phe Leu His Glu Gly Ser
370 375 380
Lys Thr Leu Val Asp Gly Leu Val Asn Ile Glu Lys Leu His Ser Val
385 390 395 400
Ala Glu Lys Val Arg Thr Ile Arg Lys Tyr Arg Ser Arg Pro Leu Cys
405 410 415
Leu Asp Met Glu Ala Ser Pro Asn His Leu Gln Thr Lys Ala Tyr Val
420 425 430
Arg Gln Phe Gln Val Ile Asp Asn Gln Asn Leu Leu Phe Glu Leu Ser
435 440 445
Tyr Lys Leu Glu Ala Asn Ser Gln
450 455

<210> 98

<211> 715

<212> PRT

<213> Homo sapiens

<400> 98
Met Ser Gln Val Met Ser Ser Pro Leu Leu Ala Gly Gly His Ala Val
1 5 10 15

Ser Leu Ala Pro Cys Asp Glu Pro Arg Arg Thr Leu His Pro Ala Pro
20 25 30

Ser Pro Ser Leu Pro Pro Gln Cys Ser Tyr Tyr Thr Thr Glu Gly Trp
35 40 45

Gly Ala Gln Ala Leu Met Ala Pro Val Pro Cys Met Gly Pro Pro Gly
50 55 60

Arg Leu Gln Gln Ala Pro Gln Val Glu Ala Lys Ala Thr Cys Phe Leu
65 70 75 80

Pro Ser Pro Gly Glu Lys Ala Leu Gly Thr Pro Glu Asp Leu Asp Ser
85 90 95

Tyr Ile Asp Phe Ser Leu Glu Ser Leu Asn Gln Met Ile Leu Glu Leu
100 105 110

Asp Pro Thr Phe Gln Leu Leu Pro Pro Gly Thr Gly Gly Ser Gln Ala
115 120 125

Glu Leu Ala Gln Ser Thr Met Ser Met Arg Lys Lys Glu Glu Ser Glu
130 135 140

Ala Leu Asp Ile Lys Tyr Ile Glu Val Thr Ser Ala Arg Ser Arg Cys
145 150 155 160

His Asp Trp Pro Gln His Cys Ser Ser Pro Ser Val Thr Pro Pro Phe
165 170 175

Gly Ser Pro Arg Ser Gly Gly Leu Leu Leu Ser Arg Asp Val Pro Arg
180 185 190

Glu Thr Arg Ser Ser Ser Glu Ser Leu Ile Phe Ser Gly Asn Gln Gly
195 200 205

Arg Gly His Gln Arg Pro Leu Pro Pro Ser Glu Gly Leu Ser Pro Arg
210 215 220

Pro Pro Asn Ser Pro Ser Ile Ser Ile Pro Cys Met Gly Ser Lys Ala
225 230 235 240

Ser Ser Pro His Gly Leu Gly Ser Pro Leu Val Ala Ser Pro Arg Leu
245 250 255

Glu Lys Arg Leu Gly Gly Leu Ala Pro Gln Arg Gly Ser Arg Ile Ser
260 265 270

Val Leu Ser Ala Ser Pro Val Ser Asp Val Ser Tyr Met Phe Gly Ser
275 280 285

Ser Gln Ser Leu Leu His Ser Ser Asn Ser Ser His Gln Ser Ser Ser
290 295 300

Arg Ser Leu Glu Ser Pro Ala Asn Ser Ser Ser Ser Leu His Ser Leu
305 310 315 320

Gly Ser Val Ser Leu Cys Thr Arg Pro Ser Asp Phe Gln Ala Pro Arg
325 330 335

Asn Pro Thr Leu Thr Met Gly Gln Pro Arg Thr Pro His Ser Pro Pro
340 345 350

Leu Ala Lys Glu His Ala Ser Ile Cys Pro Pro Ser Ile Thr Asn Ser
355 360 365

Met Val Asp Ile Pro Ile Val Leu Ile Asn Gly Cys Pro Glu Pro Gly
370 375 380

Ser Ser Pro Pro Gln Arg Thr Pro Gly His Gln Asn Ser Val Gln Pro
385 390 395 400

Gly Ala Ala Ser Pro Ser Asn Pro Cys Pro Ala Thr Arg Ser Asn Ser
405 410 415

Gln Thr Leu Ser Asp Ala Pro Phe Thr Thr Cys Pro Glu Gly Pro Ala
420 425 430

Arg Asp Met Gln Pro Thr Met Lys Phe Val Met Asp Thr Ser Lys Tyr
435 440 445

Trp Phe Lys Pro Asn Ile Thr Arg Glu Gln Ala Ile Glu Leu Leu Arg
450 455 460

Lys Glu Glu Pro Gly Ala Phe Val Ile Arg Asp Ser Ser Ser Tyr Arg
465 470 475 480

Gly Ser Phe Gly Leu Ala Leu Lys Val Gln Glu Val Pro Ala Ser Ala
485 490 495

Gln Asn Arg Pro Gly Glu Asp Ser Asn Asp Leu Ile Arg His Phe Leu
500 505 510

Ile Glu Ser Ser Ala Lys Gly Val His Leu Lys Gly Ala Asp Glu Glu
515 520 525

Pro Tyr Phe Gly Ser Leu Ser Ala Phe Val Cys Gln His Ser Ile Met
530 535 540

Ala Leu Ala Leu Pro Cys Lys Leu Thr Ile Pro Gln Arg Glu Leu Gly
545 550 555 560
Gly Ala Asp Gly Ala Ser Asp Ser Thr Asp Ser Pro Ala Ser Cys Gln
565 570 575
Lys Lys Ser Ala Gly Cys His Thr Leu Tyr Leu Ser Ser Val Ser Val
580 585 590
Glu Thr Leu Thr Gly Ala Leu Ala Val Gln Lys Ala Ile Ser Thr Thr
595 600 605
Phe Glu Arg Asp Ile Leu Pro Thr Pro Thr Val Val His Phe Glu Val
610 615 620
Thr Glu Gln Gly Ile Thr Leu Thr Asp Val Gln Arg Lys Val Phe Phe
625 630 635 640
Arg Arg His Tyr Pro Leu Thr Thr Leu Arg Phe Cys Gly Met Asp Pro
645 650 655
Glu Gln Arg Lys Trp Gln Lys Tyr Cys Lys Pro Ser Trp Ile Phe Gly
660 665 670
Phe Val Ala Lys Ser Gln Thr Glu Pro Gln Glu Asn Val Cys His Leu
675 680 685
Phe Ala Glu Tyr Asp Met Val Gln Pro Ala Ser Gln Val Ile Gly Leu
690 695 700
Val Thr Ala Leu Leu Gln Asp Ala Glu Arg Met
705 710 715
<210> 99

<211> 35

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 99
ccatatataa aaccactgtc ctgtcctttg tggct 35
<210> 100

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 100
cccccatctg tctgtctata tttgtc 26
<210> 101

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 101

tgcctacgct gacgactatg tg 22
<210> 102

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 102
tttggttttc tacaactgtt gctat 25
<210> 103

<211> 19

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 103
gggctccaca caccagatg 19
<210> 104

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 104
acgctctgag caccctctac a 21
<210> 105

<211> 31

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 105
tgtcacaggg actgaaaacc tctcctcatg t 31
<210> 106

<211> 17

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 106
cccaaggcca cgagctt 17
<210> 107

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 107
tgttgctctc ttaacgaatc gaaa 24
<210> 108

<211> 29

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 108
ctggtcaaac aaactctctg aacccctcc 29
<210> 109

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 109
tggtgaggaa aagcggacat 20
<210> 110

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 110
ctggcttgga ggacagtgaa g 21
<210> 111

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 111
ccaagccctc cccatcccat gtat 24
<210> 112

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 112
gaggtgtcgt accgcgttct a 21
<210> 113

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 113
ccgttctgct cttccctgtc t 21
<210> 114

<211> 23

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 114
ccagacccgc ttcactgacc tgc 23
<210> 115

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 115
cgcctgtact tcagcatgga 20
<210> 116

<211> 18

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 116
gcggttcagc tggtggaa 18
<210> 117

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 117
accccgaggc atcaccacaa atcat 25
<210> 118

<211> 23

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 118
agttctgcct ctctgacaac cat 23
<210> 119

<211> 23

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 119
taggctcaga gtcagaccca aac 23
<210> 120

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 120
ccctcgtggg cttgtgctcg g 21
<210> 121

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 121
aagccgccag ttcatctttt t 21
<210> 122

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 122
cttgtggttc aagtcaaatg ttcag 25
<210> 123

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 123

tctgcctgcg ctctcgtcgg t 21
<210> 124

<211> 18

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 124
gggctgggca cctgactt 18
<210> 125

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 125
cccaacaagg gtcccagact 20
<210> 126

<211> 17

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 126
cggcgcattg agcggcg 17
<210> 127

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 127
cccaagggac ttcgtgaatg 20
<210> 128

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 128
ggcgatccct gatgacaagt a 21
<210> 129

<211> 29

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer#
<400> 129
agcaccaact gtgaaccagg tacaatggc 29
<210> 130

<211> 19

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 130
gagggaggct ctgctttgg 19
<210> 131

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 131
tcacaactag cgggtgagga g 21
<210> 132

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 132
tgcagaggaa cggcgtgagc g 21
<210> 133

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 133
tgaggtttcc tcccaaatcg ta 22
<210> 134

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 134
cagctcaagg gaagctgtca tc 22
<210> 135

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 135
cccccacatg ttccccaaga tgct 24
<210> 136

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 136
ggaggcgcta aaggtctacg t 21
<210> 137

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 137
tgatgcttcg caggtcagta a 21
<210> 138

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 138
ctcctgcccc tcctaaagct gaagcc 26
<210> 139

<211> 17

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 139
ggacgcgtgg gcttttc 17
<210> 140

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 140
tgtggctgtg gacacctttc 20
<210> 141

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 141
ccacaagctg aaggcagaca aggcc 25
<210> 142

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 142
gcggattctc atggaacaca 20
<210> 143

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 143
ggtcagccag gagcttcttg 20
<210> 144

<211> 23

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 144
accaccttgc gcaggttgtc cag 23
<210> 145

<211> 18

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 145

cgcatgcacg acctgaac 18
<210> 146

<211> 23

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 146
gtctcgatct tggacagctt ctg 23
<210> 147

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 147
acactgtcca cacggcccga gg 22
<210> 148

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 148
ctgggcagaa tggaaggatc t 21
<210> 149

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 149
gggactctag cagacccaca ct 22
<210> 150

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 150
cacccacctg gattccctgt tc 22
<210> 151

<211> 23

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 151
ccttcagaca ggcgtagatg atg 23
<210> 152

<211> 29

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 152
gggtattatt tctttattag gtgccactt 29
<210> 153

<211> 30

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 153
ttccctaagg ctttcagtac ccaggatctg 30
<210> 154

<211> 18

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 154
ccagcttggc cctttcct 18
<210> 155

<211> 23

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 155
gaatgggtcg cttttgttct tag 23
<210> 156

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 156
tcacggacct cagcctgccc ct 22
<210> 157

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 157
tggtgaaggt gtcagccatg t 21
<210> 158

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 158
tcagagtgca gcaatggctt t 21
<210> 159

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 159
acctccttcc ccagctcccc 20
<210> 160

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 160
ggcaacatct tacttgtcct ttga 24
<210> 161

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 161
ccaaggaagc acagacaact atttc 25
<210> 162

<211> 30

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 162
tcctccctat ccatggcact aaaccacttc 30
<210> 163

<211> 19

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 163
tgggcaaggg ctcctatct 19
<210> 164

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 164
gttaccccctg gcagacgtat g 21
<210> 165

<211> 31

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 165
tgcctctgag tctgaatctc ccaaagagag a 31
<210> 166

<211> 31

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 166
gagtagttat gtgattattt cagctcttga c 31
<210> 167

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 167

tcaaatgttg tccccgagtc t 21
<210> 168

<211> 34

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 168
cagaaattcg gaagacagaa ctattgtcat gcct 34
<210> 169

<211> 27

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 169
gattagtaac ccatagcagt tgaaggt 27
<210> 170

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 170
atttactgac ggtggtctga acatac 26
<210> 171

<211> 31

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 171
tgacagactc caaatcacaa gcacagtcaa c 31
<210> 172

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 172
tgatggtttg gaggaaagtt tattt 25
<210> 173

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 173
tttggttggg tctttagagg aatc 24
<210> 174

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 174
tgccaaccat gcatcaggta gccc 24
<210> 175

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 175
cagctcacct ggcaacttca 20
<210> 176

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 176
cctgattttc ccagcgatgt 20
<210> 177

<211> 19

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 177
cgccgctccc ggttctgct 19
<210> 178

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 178
tggccaagcg taagctgatt 20
<210> 179

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> PCR primer
<400> 179
gctgcagtga tcggatcatc t 21
<210> 180

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> MLLT6
<400> 180
caccatggag cccatcgtgc tg 22
<210> 181

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> MLLT6 for
<400> 181
atccccgagg tgcaatttg 19
<210> 182

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> MLLT6 rev
<400> 182
agcgatcatg aggcacgtac t 21
<210> 183

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> ZNF144
<400> 183
cctgccagag ataggagacc cagacagct 29
<210> 184

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> ZNF144 for
<400> 184
atccccctga gccttttca 19
<210> 185

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> ZNF144 rev
<400> 185
cagcctctgg tcccaccat 19
<210> 186

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> PIP5K2B
<400> 186
tgatcatcaa ttccaaacct ctcccgaa 28
<210> 187

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> PIP5K2B for
<400> 187
ccccatggtg ttccgaaac 19
<210> 188

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> PIP5K2B rev
<400> 188
tgccaggagc ctccatacc 19
<210> 189

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> TEM7
<400> 189

cagccttcta aaacacaatg tattcatgt 29
<210> 190

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> TEM7 for
<400> 190
cctgaactta atggtagaat tcaaagatc 29
<210> 191

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> TEM7 rev
<400> 191
tattaacact gagaatccat gcagaga 27
<210> 192

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> ZNFN1A3
<400> 192
tatctggtct cagggattgc tcctatgtat tcagc 35
<210> 193

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> ZNFN1A3 for
<400> 193
cacagagccc tgctgaagtg 20
<210> 194

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> ZNFN1A3 rev
<400> 194
gcgaggtcat tggtttttag aaa 23
<210> 195

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> WIRE
<400> 195
ctgtgatccg aaatggtgcc ag 22
<210> 196

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> WIRE for
<400> 196
ccgtctccac atccaaacct 20
<210> 197

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> WIRE rev
<400> 197
acccatgcat tcggtatggt 20
<210> 198

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> PSMB3
<400> 198
agtggcacct gcgccgaaca a 21
<210> 199

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> PSMB3 for
<400> 199
ccccatggtg actgatgact t 21
<210> 200

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> PSMB3 rev
<400> 200
ccagagggac tcacacattc c 21
<210> 201

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> MGC9753
<400> 201
ccagaaactt tccatcccaa aggcagtct 29
<210> 202

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> MGC9753 for
<400> 202
ctgccccaca ggaatagaat g 21
<210> 203

<211> 23

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> MGC9753 rev
<400> 203
aaaaatccag tctgcttcaa cca 23
<210> 204

<211> 20

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> ORMDL3
<400> 204
agctgcccca gctccacgga 20
<210> 205

<211> 21

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> ORMDL3 for
<400> 205
tccctgatga gcgtgcttat c 21
<210> 206

<211> 28

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> ORMDL3 rev
<400> 206
tctcagtact tattgattcc aaaaatcc 28
<210> 207

<211> 25

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> MGC15482
<400> 207
tccagtggaa gcaaccccag tgttc 25
<210> 208

<211> 25

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> MGC15482 for
<400> 208
cacttctaga gctaccgtgg agtct 25
<210> 209

<211> 22

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> MGC15482 rev
<400> 209
ccctcacttt gtaacccttg ct 22
<210> 210

<211> 20

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> PPP1R1B
<400> 210
cagcgtggcg caacaaccca 20
<210> 211

<211> 21

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> PPP1R1B for
<400> 211

gggattgttt cgccacacat a 21
<210> 212

<211> 20

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> PPP1R1B rev
<400> 212
ccgatgttaa ggcccatagc 20
<210> 213

<211> 27

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> MGC14832
<400> 213
taaaatgtcc ggccaacatg agttccc 27
<210> 214

<211> 17

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> MGC14832 for
<400> 214
cgcagtgcct ggcacat 17
<210> 215

<211> 20

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> MGC14832 rev
<400> 215
gacaccccct gacctatgga 20
<210> 216

<211> 25

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> LOC51242
<400> 216
cagtgacctc tcccgttccc ttgga 25
<210> 217

<211> 20

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> LOC51242 for
<400> 217
tgggtccctg tgtcctcttc 20
<210> 218

<211> 20

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> LOC51242 for
<400> 218
agggtcagga gggagaaaac 20
<210> 219

<211> 26

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> FLJ20291
<400> 219
ccagtgccca cccgttaaag agtcaa 26
<210> 220

<211> 24

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> FLJ20291 for
<400> 220
ttgtgggaca ctcagtaact ttgg 24
<210> 221

<211> 20

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> FLJ20291 rev
<400> 221
acaagcactc ccaccgagat 20
<210> 222

<211> 24

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> PRO2521
<400> 222
agtctgtcct cactgccatc gcca 24
<210> 223

<211> 21

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> PRO2521 for
<400> 223
aagcctctgg gttttccctt t 21
<210> 224

<211> 20

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> PRO2521 rev
<400> 224
cccactggtg acaggatggt 20
<210> 225

<211> 23

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> LINK-GEFII
<400> 225
catctgacat ctttcccgtg gag 23
<210> 226

<211> 21

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> LINK-GEFII for
<400> 226
ctttgcacga tgtctcaacc a 21
<210> 227

<211> 18

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> LINK-GEFII rev
<400> 227
tttcccgtgg agcaggaa 18
<210> 228

<211> 26

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> CTEN
<400> 228
ccgccgccta atatgcaaca ttaggg 26
<210> 229

<211> 23

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> CTEN for
<400> 229
cgagtattcc aaagctggta tcg 23
<210> 230

<211> 24

<212> DNA

<213> ARTIFICIAL SEQUENCE

<220>

<223> CTEN rev
<400> 230
atcacagaga gatggccctt atct 24
<210> 231

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S946 forward primer
<400> 231
acagtctatc aagcagaaaa atcct 25
<210> 232

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S946 reverse primer
<400> 232
tgccgtgcca gagaga 16
<210> 233

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1181 forward primer
<400> 233

gacaacagag cgagactccc 20
<210> 234

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1181 reverse primer
<400> 234
gcccagcctg tcacttattc 20
<210> 235

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2026 forward primer
<400> 235
tggtcattcg acaacgaa 18
<210> 236

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2026 reverse primer
<400> 236
cagcattgga tgcaatcc 18
<210> 237

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S838 forward primer
<400> 237
ctccagaatc cagaccatga 20
<210> 238

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S838 reverse primer
<400> 238
aggacagtgt gtagcccttc 20
<210> 239

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S250 forward primer
<400> 239
ggaagaatca aatagacaat 20
<210> 240

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S250 reverse primer
<400> 240
gctggccata tatatattta aacc 24
<210> 241

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1818 forward primer
<400> 241
cataggtatg ttcagaaatg tga 23
<210> 242

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1818 reverse primer
<400> 242
tgcctactgg aaaccaga 18
<210> 243

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S614 forward primer
<400> 243
aaggggaagg ggctttcaaa gct 23
<210> 244

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S614 reverse primer
<220>

<221> misc_feature

<222> (1)..(1)

<223> n=a, c, g or t

<400> 244
nggaggttgc agtgagccaa gat 23
<210> 245

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2019 forward primer
<400> 245
caaaagctta tgatgctcaa acc 23
<210> 246

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2019 reverse primer
<400> 246
ttgtttccct ttgactttct ga 22
<210> 247

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S608 forward primer
<400> 247
taggttcacc tctcattttc ttcag 25
<210> 248

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S608 reverse primer
<220>

<221> misc_feature

<222> (17)..(17)

<223> n=a, c, g or t

<400> 248
gtctgggtct ttatggngct tgtg 24
<210> 249

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1655 forward primer
<400> 249
cggaccagag tgttccatgg 20
<210> 250

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1655 reverse primer
<400> 250
gcatacagca ccctctacct 20
<210> 251

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2147 forward primer
<400> 251
aggggagaat aaataaaatc tgtgg 25
<210> 252

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2147 reverse primer
<400> 252
caggagtgag acactctcca tg 22
<210> 253

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S754 forward primer
<400> 253
tggattcact gactcagcct gc 22
<210> 254

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S754 reverse primer
<400> 254
gcgtgtctgt ctccatgtgt gc 22
<210> 255

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1814 forward primer
<400> 255
tccccaatga cggtgatg 18
<210> 256

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1814 reverse primer
<400> 256
ctggaggttg gcttgtggat 20
<210> 257

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2007 forward primer
<400> 257
ggtcccacga atttgctg 18
<210> 258

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2007 reverse primer
<400> 258
ccacccagaa aaacaggaga 20
<210> 259

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1246 forward primer
<400> 259
tcgatctcct gaccttgtga 20
<210> 260

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1246 reverse primer
<400> 260
ttgtcacccc attgcctttc 20
<210> 261

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1979 forward primer
<400> 261
ccttggatag attcagctcc c 21
<210> 262

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1979 reverse primer
<400> 262
cttgtccctt ctcaatcctc c 21
<210> 263

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1984 forward primer
<400> 263
ttaagcaagg ttttaattaa gctgc 25
<210> 264

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1984 reverse primer
<400> 264
gattacagtg ctccctctcc c 21
<210> 265

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> G11580 forward primer
<400> 265
ggttttaatt aagctgcatg gc 22
<210> 266

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> G11580 reverse primer
<400> 266
gattacagtg ctccctctcc c 21
<210> 267

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1867 forward primer
<400> 267

agtttgacac tgaggctttg 20

<210> 268

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1867 reverse primer

<400> 268

tttagacttg gtaactgccg 20

<210> 269

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1788 forward primer

<400> 269

tgcagatgcc taagaacttt tcag 24

<210> 270

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1788 reverse primer

<400> 270

gccatgatct cccaaagcc 19

<210> 271

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1836 forward primer

<400> 271

tcgaggttat ggtgagcc 18

<210> 272

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1836 reverse primer
<400> 272
aaactgtgtg tgtcaaagga tact 24
<210> 273

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1787 forward primer
<400> 273
gctgatctga agccaatga 19
<210> 274

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1787 reverse primer
<400> 274
tacatgaagg catggtctg 19
<210> 275

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1660 forward primer
<400> 275
ctaatataat cctgggcaca tgg 23
<210> 276

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1660 reverse primer
<400> 276
gctgcggacc agacagat 18
<210> 277

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2154 forward primer
<400> 277
gataaaaaca agcactggct cc 22
<210> 278

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2154 reverse primer
<400> 278
cccacggctt tcttgatcta 20
<210> 279

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1955 forward primer
<400> 279
tgtaatgtaa gccccatgag g 21
<210> 280

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1955 reverse primer
<400> 280
cactcaactc aacagtctaa aggtg 25
<210> 281

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2098 forward primer
<400> 281
gtgagttcaa gcatagtaat tatcc 25
<210> 282

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2098 reverse primer
<400> 282
attcagcctc agttcactgc ttc 23
<210> 283

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S518 forward primer
<400> 283
gatccagtgg agactcagag 20
<210> 284

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S518 reverse primer
<400> 284
tagtctctgg gacacccaga 20
<210> 285

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S518 forward primer
<400> 285
attcctgagt gtctaccctg ttgag 25
<210> 286

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S518 reverse primer
<400> 286
actgactgcg ccactgc 17
<210> 287

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D11S4358 forward primer
<400> 287
tcgagaagga caaaatcacc 20
<210> 288

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D11S4358 reverse primer
<400> 288
gaacagggtt agtccattcg 20
<210> 289

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S964 forward primer
<400> 289

gttctttcct cttgtgggg 19
<210> 290

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S964 reverse primer
<400> 290
agtcagctga gattgtgcc 19
<210> 291

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D19S1091 forward primer
<400> 291
caagccaaga catcccagtt 20
<210> 292

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D19S1091 reverse primer
<400> 292
ccccacacac agctcatatg 20
<210> 293

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1179 forward primer
<400> 293
ttttctctct cattccattg gg 22
<210> 294

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1179 reverse primer
<400> 294
gcaacagagg gagactccaa 20
<210> 295

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> D10S2160 forward primer
<400> 295
tcccatcccg taagacctc 19
<210> 296

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> D10S2160 reverse primer
<400> 296
tatggagtac ctactctatg ccagg 25
<210> 297

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1230 forward primer
<400> 297
attcaaagct ggatcccttt 20
<210> 298

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1230 reverse primer
<400> 298
agctgtgaca aatgcctgta 20
<210> 299

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1338 forward primer
<400> 299
tcacctgaga ttgggagacc 20
<210> 300

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1338 reverse primer
<400> 300
aagatggggc aggaatgg 18
<210> 301

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2011 forward primer
<400> 301
tcactgtcct ccaagccag 19
<210> 302

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2011 reverse primer
<400> 302
aaacaccaca ctctcccctg 20
<210> 303

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2011 forward primer
<400> 303
ttcttgggct tcccgtagcc 20
<210> 304

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2011 reverse primer
<400> 304
ggggcagacg acttctcctt 20
<210> 305

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2038 forward primer
<400> 305
ggggatacaa cctttaaagt tcc 23
<210> 306

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2038 reverse primer
<400> 306
attcacctaa tgaggattct tcttt 25
<210> 307

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2091 forward primer
<400> 307
gctgaaatag ccatcttgag ctac 24
<210> 308

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S2091 reverse primer
<400> 308
tccgcatcct ttttaagagg cac 23
<210> 309

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S649 forward primer
<400> 309
ctttcactct ttcagctgaa gagg 24
<210> 310

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S649 reverse primer
<400> 310
tgacgtgcta tttcctgttt tgtct 25
<210> 311

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1190 forward primer
<400> 311

gtttgttgct atgcctgc 18
<210> 312

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> D17S1190 reverse primer
<400> 312
caacacacta ccccagga 18
<210> 313

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> M87506 forward primer
<400> 313
actcctcatc tgtagggtct 20
<210> 314

<211> 20

<212> DNA

<213> Artificial Sequence

<400> 314
gagtccgcta cctgagtgct 20
<210> 315

<211> 4617

<212> DNA

<213> Homo sapiens

<400> 315
gcctggagac taacagctgc tcggaagagg agctcagcag cccgggtcgc ggaggaggag 60
ggggcggccg gcttctgctg cagccccccag gccctgaatt acctccggtg cccttcccgc 120
tgcaggactt ggtccctctg ggcgcctga gtagagggga gcagcagcag cagcagcagc 180
agcaacctcc cccgccccccg cctcctcccg ggcccctccg gccactcgcg ggtccttctc 240
ggaagggctc cttcaaaatc cgcctcagtc gcctctttcg caccaagagc tgcaacggtg 300
gctccggcgg tggggatggg accggcaaga ggccttctgg agagctggct gcttcagctg 360
cgagcctgac agacatggga ggctctgcgg gccgggagct ggacgcgggg aggaaaccca 420
agttgacaag aactcaaagt gccttttctc cggtctcctt cagccccctg ttcacaggtg 480
aaactgtgtc gcttgtggat gtggacattt ctcagcgggg cctgacctct ccacacccctc 540
caactccccc tcctcctccg agaagaagcc tcagcctcct agatgatatc agtgggacgc 600
tgcctacatc tgtccttgtg gctccgatgg ggtcttcctt gcagtctttc cccctacctc 660
cgcctcctcc accccatgcc ccagatgcat ttccccggat tgctcccatc cgagcagctg 720

```
aatccctgca cagccaaccc ccacagcacc tccagtgtcc cctctaccgg cctgactcga
gcagctttgc agccagcctt cgagagttgg agaagtgtgg ttggtattgg gggccaatga     780
attgggaaga tgcagagatg aagctgaaag ggaaaccaga tggttctttc ctggtacgag     840
acagttctga tcctcgttac atcctgagcc tcagtttccg atcacagggt atcacccacc     900
acactagaat ggagcactac agaggaacct tcagcctgtg gtgtcatccc aagtttgagg     960
accgctgtca atctgttgta gagtttatta agagagccat tatgcactcc aagaatggaa    1020
agtttctcta tttcttaaga tccagggttc caggactgcc accaactcct gtccagctgc    1080
tctatccagt gtcccgattc agcaatgtca aatccctcca gcacctttgc agattccgga    1140
tacgacagct cgtcaggata gatcacatcc cagatctccc actgcctaaa cctctgatct    1200
cttatatccg aaagttctac tactatgatc ctcaggaaga ggtatacctg tctctaaagg    1260
aagcgcagct catttccaaa cagaagcaag aggtggaacc ctccacgtag cgaggggctc    1320
cctgctggtc accaccaagg gcatttggtt gccaagctcc agctttgaag aaccaaatta    1380
agctaccatg aaaagaagag gaaaagtgag ggaacaggaa ggttgggatt ctctgtgcag    1440
agactttggt tccccacgca gccctggggc ttggaagaag cacatgaccg tactctgcgt    1500
ggggctccac ctcacaccca cccctgggca tcttaggact ggagggggctc cttggaaaac    1560
tggaagaagt ctcaacactg tttctttttc agaagttttg tttttgatat ttatattact    1620
tggtatggaa aactcacctt gaaggcagtt ggggtttgtg cccgttggat tgaaagtggt    1680
gcgaagggtg agcaggtcca aagaaggggt gggaggaggg aacaggggac ggccattcag    1740
ctggtgccaa aggcagagtt agagtctgtg ctgtgggcct ggaagatggg aggaggggct    1800
tgaggtttgc aaaggactgg gagttcctga ggaaggggga atctgcctct tgttgccatg    1860
ggcagctgtt gtaggagcag gagagaaagg agggtgggtg gtctcgaaaa gaatattggg    1920
caaaacctag ccaattggcc ttagctggga gaagtagtga ctcctgcatc ctttttaag    1980
gtttaggaac ctgagttcag aaacacctct catggaagct gtactagttg tgatttactt    2040
aattccttaa gttccatgac ctgaagttaa ccccgttctt cctctgctct caacccattg    2100
cccc ttgaga taactgtaca tgtcactctg atcatggtaa cagcatccct attgcttctg    2160
ccagctgtca tggcaatcgt gtttcccatc acctgggcgg ttcagagcca gtcatgggct    2220
gctgaattta atggagcatg tttccaggtt cttcatggca aactgtactc atgacttagg    2280
agtgagtgtt acttccatgt gcctgtcagc ttgtgagggg gaatgtggag gaaggtgaga    2340
aatacagctc ccacagttgt gctcttccta gaggaagctc tcagaacgca gccctcacgg    2400
gatttcctta ggtcagagga gagcatcgca tctcacgttt ttaggtttat cactgccatc    2460
ccacttctgg gatgggaggt agcaagggct tctgtatttt cttgtgttca ttctagcaac    2520
ccagacattt ccggatcaga tcctgctggt ctccactcac tggaaagtct gccagatgcc    2580
gatttgagag ctgcctgtcc ctgctttcag gaggagcggg gagaaaaact ccaatggtct    2640
ttaatggttt ctgcagctgg ccatggccaa ttcatatgac attgtgagtt tgctttctta    2700
tagagctgct ctgggggagag gtttgctatt gagatgtaac agtggagctg ttgggtcttc    2760
atgactcctt tgcgtgtgtt ccatgggact ctctttctgg gttccccatg cttatagttg    2820
cctcgtgtca caagacagat actaatgtca ggtttgtggc ttcctgatgg tttgggtggg    2880
gccccagtgt cctggtaatt tataggactg cctcatctgg gagcattgcc ttcttcctta    2940
gtcccacgtg gagtgaccag tcttcctcct tgtagctgaa cagggaggaa acttgcacca    3000
ttacctgact gtggaagggt ggcccacaag atgagctgtg caccataaac acagcccacc    3060
tctgatttgt catgtggtac ctcttctttc cttggcttcc atggtagtat taccaactaa    3120
gcaagattgt gatcccagaa attggcttag catgtgagtg ttgcctcgtg agagtacaag    3180
taatataact cgccatcttg caggaagtgc caccccaata tagagcctga agttggaatc    3240
tgttgagatc cttgggtggc tgatatacag cctgggatct ttcttttttt tgttcctttt    3300
caaccaccca taattttaat attatttttt agtgtgtgtg tgcctggctt tgcgctagat    3360
attgtagaaa acaaaaaagg taaaagacgt aatatgtggc ctaagggagc ttttaggtga    3420
ctgctgcaca tcaagcagaa aatcaaggac tatctaaaga cgtttatagt agataagatc    3480
agggtagacc agatggtctg ggaaagttct gtgcctctga ggctttgggt tgtagtcaat    3540
ggcaggacag acagtgagat gaaaaacaca tgagcaaaag caaggaagca gaaatctgca    3600
tggcatgtac tgaacagtgc acagccctgt tagagcaaca tggttaaaga atcctttcca    3660
gtgcggtttt ctagatggaa gcttcccagc caccaggcag acctgagtgc cgaggggtta    3720
tgatggtgag gtggagccca aagcccaaag gagtcagcaa ggctcctgcc cattgccagg    3780
gcctcactat ggtcagctca ggccatgtga gggaggcaga gcctccgcac cccctgtgtt    3840
actggggttt cttctggaga actcatacat tcaggtacaa aacaaaccaa ctgaggaggt    3900
gtgacccaac ctcaccaccc acctttctcc tcctggggag tgtcgtgttg aactgtgtct    3960
gtgtcagtgc actggtccca gccctggccg cagcctagtc ctttctctgt ggagtgggct    4020
gcaaaacagc agcgcaagag gagacagtct tcgtcaggtg tgggtgtttg gtgcctgcat    4080
gggtggcctt gcacagagca gttaggggaa gatgaggggg aggcatgggg ctgggccagc    4140
tctctgggat acagcctggt cagagagtaa atggagcaga ggaacaggtg tggcagcacc    4200
tgcccttcac ctccctgacc aggcccgtcc ctccttcgcc agtgctgtgc agagctcatt    4260
taaatgtatt cctttctagg tctgggcgcg gtggctcacg cctgtaatcc cagcactttg    4320
ggaggctgag gtgggtggat catgaggtcg ggagttcaag accagcttgg ccaaacagtg    4380
aaaccctatc tctactaaaa tacaaaaatt agctgggcat ggtggcatgc gcctgtagtc    4440
ccagctactt gggaggctga gacaggagaa ttgcttgaac ccaggaggcg gaggttgcac    4500
tgagccgaga ttgcgccact gccctctagc ctagacgaca gagtgagact ccatctc        4560
                                                                     4617
<210>  316
```

<211> 1713

<212> DNA

<213> Homo sapiens

<400> 316
gagcacgtgc gcgtcctcgc tgctcacatt tcggcggagg gcgcgctccc tggaaaattc 60
cactcctgtg ctagctccac cctatgcggc ttttctccta cccgacgctc ttcactctca 120
gctcccttcc cggcggcctt tgcgggaaca agatggcagc cccatacct caagggttct 180
cttgtttatc gagggttttg ggctggtggt ctcggcagcc agttctggtg actcagtccg 240
cagctatagt tccagtaaga actaaaaaac gtttcacacc tcctatttat caacctaaat 300
ttaaaacaga aaaggagttt atgcaacatg cccggaaagc aggattggtt attcctccag 360
aaaaatcgga ccgttccata catctggcct gtacagctgg tatatttgat gcctatgttc 420
ctcctgaggg tgatgcacgc atatcatctc tttcaaagga gggactgata gagagaactg 480
aacgaatgaa gaagactatg gcatcacaag tgtcaatccg gaggataaaa gactatgatg 540
ccaactttaa aataaaggac ttccctgaaa aagctaagga tatctttatt gaagctcacc 600
tttgtctaaa taactcagac catgaccgac ttcatacctt ggtaactgaa cactgttttc 660
cagacatgac ttgggacatc aaatataaga ccgtccgctg gagctttgtg gaatctttag 720
agccctctca tgttgttcaa gttcgctgtt caagtatgat gaaccagggc aacgtgtacg 780
gccagatcac cgtacgcatg cacacccggc agactctggc catctatgac cggtttggcc 840
ggttgatgta tggacaggaa gatgtaccca aggatgtcct ggagtatgtt gtattcgaaa 900
agcagttgac aaacccctat ggaagctgga gaatgcatac caagatcgtt cccccatggg 960
cacccccctaa gcagcccatc cttaagacgg tgatgatccc tggccctcag ctgaaaccag 1020
aagaagaata tgaagaggca caaggagagg cccagaagcc tcagctagcc tgatgacaaa 1080
aatgacttct agggtgaagc ctgggtgatg aggctgctgg aagctttgaa gtctcccatt 1140
cccctcatgc tataaaaaga actacctttg ttctctccca tcctgctcag gtcttttcag 1200
cagtctcatc atcagcaacc atgactgatg actgggccct agcaggtggc aggtataaca 1260
tggccatgga cactcttctt ttttaaattt tatgtctagc ttctgagtct agatgaaaga 1320
cagtatgttt cagagaacat tggatatcag tttttcccac agcagggact gtgagagaca 1380
accagcagca tcctctttgt aatcacaggg cagggatcag agtttgaaat gaaatgttgt 1440
cagggtgttg gaaaaatttt ggtgagttct gcacatttcc cctggttcag gctgggcatg 1500
gaccagcctt cagatggcag aagtggaaga tgagcctact tgtgagcgat gtgactttaa 1560
ggaaatgaag actggggaag aataattagt gtttataaga catttaagag gccctttttc 1620
atatactgac tcactgatga atcagcattt gcattttatg gaaaaatata aatccaaaga 1680
aataatttat cccttaaaaa aaaaaaaaaa aaa 1713

<210> 317

<211> 5632

<212> DNA

<213> Homo sapiens

<400> 317
gtggcgatca tggcaagtta gaagttttct gactcctttc ggaggagcct ccgggacccc 60
ggggagtaac aggtgtctgg aggctgaagg gtggaggggt tcctggattt gggttttgct 120
tgtgaaactc ccctccaccc tcctctctcg cacccaccca ccccctcacc cccttctttt 180
tccgtccttg gaaaatggtg tccaagctca cgtcgctcca gcaagaactc ctgagcgccc 240
tgctgagctc cggggtcacc aaggaggtgc tggttcaggc cttggaggag ttgctgccat 300
ccccgaactt cggggtgaag ctggagacgc tgcccctgtc ccctggcagc ggggccgagc 360
ccgacaccaa gccggtcttc catactctca ccaacggcca cgccaagggc cgcttgtccg 420
gcgacgaggg ctccgaggac ggcgacgact atgacacacc tcccatcctc aaggagctgc 480
aggcgctcaa caccgaggag gcggcggagc agcgggcgga ggtggaccgg atgctcagtg 540
aggacccttg gagggctgct aaaatgatca agggttacat gcagcaacac aacatccccc 600
agagggaggt ggtcgatgtc accggcctga accagtcgca cctctcccag catctcaaca 660
agggcacccc tatgaagacc cagaagcgtg ccgctctgta cacctggtac gtcagaaagc 720
aacgagagat cctccgacaa ttcaaccaga cagtccagag ttctggaaat atgacagaca 780
aaagcagtca ggatcagctg ctgtttctct ttccagagtt cagtcaacag agccatgggc 840
ctgggcagtc cgatgatgcc tgctctgagc ccaccaacaa gaagatgcgc cgcaaccggt 900
tcaaatgggg gcccgcgtcc cagcaaatct tgtaccaggc ctacgatcgg caaaagaacc 960

ccagcaagga agagagagag gccttagtgg aggaatgcaa cagggcagaa tgtttgcagc
gaggggtgtc cccctccaaa gcccacggcc tgggctccaa cttggtcact gaggtccgtg 1020
tctacaactg gtttgcaaac cgcaggaagg aggaggcatt ccggcaaaag ctggccatgg 1080
acgcctatag ctccaaccag actcacagcc tgaaccctct gctctcccac ggctcccccc 1140
accaccagcc cagctcctct cctccaaaca agctgtcagg agtgcgctac agccagcagg 1200
gaaacaatga gatcacttcc tcctcaacaa tcagtcacca tggcaacagc gccatggtga 1260
ccagccagtc ggttttacag caagtctccc cagccagcct ggacccaggc cacaatctcc 1320
tctcacctga tggtaaaatg atctcagtct caggaggagg tttgccccca gtcagcacct 1380
tgacgaatat ccacagcctc tcccaccata atccccagca atctcaaaac ctcatcatga 1440
cacccctctc tggagtcatg gcaattgcac aaagcctcaa cacctcccaa gcacagagtg 1500
tccctgtcat caacagtgtg gccggcagcc tggcagccct gcagcccgtc cagttctccc 1560
agcagctgca cagccctcac cagcagcccc tcatgcagca gagcccaggc agccacatgg 1620
cccagcagcc cttcatggca gctgtgactc agctgcagaa ctcacacatg tacgcacaca 1680
agcaggaacc cccccagtat tcccacacct cccggtttcc atctgcaatg gtggtcacag 1740
ataccagcag catcagtaca ctcaccaaca tgtcttcaag taaacagtgt cctctacaag 1800
cctggtgatg cccacacacc acttacttcg tgcgcaacaa caaggaccct gttttccaca 1860
ccatcaccct ctgggcagct gtcatggaaa agcccagtga cctgaccagc acctgcgaga 1920
ggtccctgct tacctgacgg acgtcctgct ggcacctcag acaatccact ctcaggagcg 1980
cagcccgaag cccagtttcc cttctatgca gtattgccac aatgcctctc ccacgatgtc 2040
aaggactcct gtctgtcctg gaggtgggag acaaggaacc tccgaagagg aagcaagaaa 2100
gccgtactgt ctatgttgtg atccttcatc gaacaaactg atgcgaaaac ttgaatctgt 2160
tactgaaatg aggagagaag gacatgtgct attgaactga gccaaacaca ctgtaaatat 2220
ccacagactc cctcccctgc ccccatccca aatgatcttg agatttcttt taaagaagta 2280
aatttgtcca atggctgtaa actataaact actgtaatta agtgcaattt cccctctgtg 2340
tcctctcccc tctgccctgt atataatact aaagtgtcta ttagttttct ttgtaaaggt 2400
cagagtcaaa atttcaaaag tgatctgtcc cctctcccct catggagaaa catcctaagt 2460
gggaagtgaa gccccttgtc ctctcccgcg aggcctggac acttatgggg acagcatacc 2520
ttggactgac taccagctaa ctccagtctc ctgacattaa gacacacctc tggatccctg 2580
gaggggctga atgtagtgtg tcagagtaac atgccagctt cctgtgggcc aggagctcag 2640
ccgtgcactc cctaagaaac cccagggcag ggaaactggc tgtttgatag cagaagaaaa 2700
agttgcagtc tcagaaagcc ttccattaaa acaatttatt ttatcactaa aaaaaagtgg 2760
cgatcatggc aagttagaag ttttctgact cctttcggag gagcctccgg gaccccgggg 2820
agtaacaggt gtctggaggc tgaagggtgg aggggttcct ggatttgggt tttgcttgtg 2880
aaactcccct ccaccctcct ctctcgcacc cacccacccc ctcacccccct tctttttccg 2940
tccttggaaa atggtgtcca agctcacgtc gctccagcaa gaactcctga gcgccctgct 3000
gagctccggg gtcaccaagg aggtgctggt tcaggccttg gaggagttgc tgccatcccc 3060
gaacttcggg gtgaagctgg agacgctgcc cctgtcccct ggcagcgggg ccgagcccga 3120
caccaagccg gtcttccata ctctcaccaa cggccacgcc aagggccgct tgtccggcga 3180
cgagggctcc gaggacggcg acgactatga cacacctccc atcctcaagg agctgcaggc 3240
gctcaacacc gaggaggcgg cggagcagcg ggcggaggtg gaccggatgc tcagtgagga 3300
cccttggagg gctgctaaaa tgatcaaggg ttacatgcag caacacaaca tcccccagag 3360
ggaggtggtc gatgtcaccg gcctgaacca gtcgcacctc tcccagcatc tcaacaaggg 3420
cacccctatg aagacccaga agcgtgccgc tctgtacacc tggtacgtca gaaagcaacg 3480
agagatcctc cgacaattca accagacagt ccagagttct ggaaatatga cagacaaaag 3540
cagtcaggat cagctgctgt ttctctttcc agagttcagt caacagagcc atgggcctgg 3600
gcagtccgat gatgcctgct ctgagcccac caacaagaag atgcgccgca accggttcaa 3660
atgggggccc gcgtcccagc aaatcttgta ccaggcctac gatcggcaaa agaaccccag 3720
caaggaagag agagaggcct tagtggagga atgcaacagg gcagaatgtt tgcagcgagg 3780
ggtgtccccc tccaaagccc acggcctggg ctccaacttg gtcactgagg tccgtgtcta 3840
caactggttt gcaaaccgca ggaaggagga ggcattccgg caaaagctgg ccatggacgc 3900
ctatagctcc aaccagactc acagcctgaa ccctctgctc tcccacggct cccccccacca 3960
ccagcccagc tcctctcctc caaacaagct gtcaggagtg cgctacagcc agcagggaaa 4020
caatgagatc acttcctcct caacaatcag tcaccatggc aacagcgcca tggtgaccag 4080
ccagtcggtt ttacagcaag tctccccagc cagcctggac ccaggccaca atctcctctc 4140
acctgatggt aaaatgatct cagtctcagg aggaggtttg cccccagtca gcaccttgac 4200
gaatatccac agcctctccc accataatcc ccagcaatct caaaacctca tcatgacacc 4260
cctctctgga gtcatggcaa ttgcacaaag cctcaacacc tcccaagcac agagtgtccc 4320
tgtcatcaac agtgtggccg gcagcctggc agccctgcag cccgtccagt tctcccagca 4380
gctgcacagc cctcaccagc agcccctcat gcagcagagc ccaggcagcc acatggccca 4440
gcagcccttc atggcagctg tgactcagct gcagaactca cacatgtacg cacacaagca 4500
ggaacccccc cagtattccc acacctcccg gtttccatct gcaatggtgg tcacagatac 4560
cagcagcatc agtacactca ccaacatgtc ttcaagtaaa cagtgtcctc tacaagcctg 4620
gtgatgccca cacaccactt acttcgtgcg caacaacaag gaccctgttt tccacaccat 4680
caccctctgg gcagctgtca tggaaaagcc cagtgacctg accagcacct gcgagaggtc 4740
cctgcttacc tgacggacgt cctgctggca cctcagacaa tccactctca ggagcgcagc 4800
ccgaagccca gtttcccttc tatgcagtat tgccacaatg cctctcccac gatgtcaagg 4860
4920

actcctgtct gtcctggagg tgggagacaa ggaacctccg aagaggaagc aagaaagccg 4980
tactgtctat gttgtgatcc ttcatcgaac aaactgatgc gaaaacttga atctgttact 5040
gaaatgagga gagaaggaca tgtgctattg aactgagcca aacacactgt aaatatccac 5100
agactccctc ccctgccccc atcccaaatg atcttgagat ttcttttaaa gaagtaaatt 5160
tgtccaatgg ctgtaaacta taaactactg taattaagtg caatttcccc tctgtgtcct 5220
ctcccctctg ccctgtatat aatactaaag tgtctattag ttttctttgt aaaggtcaga 5280
gtcaaaattt caaaagtgat ctgtcccctc tcccctcatg gagaaacatc ctaagtggga 5340
agtgaagccc cttgtcctct cccgcgaggc ctggacactt atggggacag catacettgg 5400
actgactacc agctaactcc agtctcctga cattaagaca cacctctgga tccctggagg 5460
ggctgaatgt agtgtgtcag agtaacatgc cagcttcctg tgggccagga gctcagccgt 5520
gcactcccta agaaacccca gggcagggaa actggctgtt tgatagcaga agaaaaagtt 5580
gcagtctcag aaagccttcc attaaaacaa tttattttat cactaaaaaa aa 5632

<210> 318

<211> 3123

<212> DNA

<213> Homo sapiens

<400> 318
gaactgtggc gctttctggg taaagatgga cgtccacgat ctctttcgcc ggctcggcgc 60
gggggccaaa ttcgacacga gacgcttctc ggcagacgca gctcgattcc agataggaaa 120
aaggaaatat gactttgatt cttcggaggt gcttcaggga ctggactttt ttggaaacaa 180
gaagtctgtc ccaggtgtgt gtggagcatc acaaacacat cagaagcccc aaaatggaga 240
gaaaaaagaa gagagcctaa ctgaaaggaa gagggagcag agcaagaaaa aaaggaagac 300
gatgacttca gagacagggt ttcaccatgt tggccagtat ggtctcgatc tcctgacctc 360
gtgattcacc caccttgggc tcccaaagtg ctgggattac agatgtgagc caccacgccc 420
agccagaaat tgcttcccaa gaagaaggtg ctactataca gtggatgtca tctgtagaag 480
caaagattga agacaaaaaa gttcagagag aaagtaaact aacttccgga aagttggaga 540
atctcagaaa agaaaagata aacttcttgc ggaataaaca caaaattcac gtccaaggaa 600
ccgatcttcc tgacccaatt gctacatttc agcaacttga ccaggaatat aaaatcaatt 660
ctcgactact tcagaacatt ctagatgcag gcttccaaat gcctacgcca atccaaatgc 720
aagccatccc agttatgctg catggtcggg aacttctggc ttctgctcca actggatctg 780
gaaaaacatt agcttttagc attcctattt taatgcagct gaaacaaccc gcaaataaag 840
gcttcagagc cctgattata tcaccaacac gagaacttgc cagccagatt cacagagagt 900
taataaaaat ttctgaggga acaggattca gaatacacat gatccacaaa gcagcagtgg 960
cagccaagaa atttggacct aaatcatcta aaagtttgga tattcttgtg actactccaa 1020
atcgactaat ctatttatta aagcaagatc cccccggaat cgacctagca agtgttgagt 1080
ggcttgtagt agacgaatca gataaactgt ttgaagatgg caaaactggg ttcagagacc 1140
agctggcttc cattttcctg gcctgcacat cccacaaggt ccgaagagct atgttcagtg 1200
caacttttgc atatgatgtt gaacagtggt gcaaactcaa cctggacaat gtcatcagtg 1260
tgtccattgg agcaaggaat tctgcagtag aaactgtaga acaagagctt ctctttgttg 1320
gatctgagac cggaaaactt ctggccgtga gagaacttgt taaaaagggt ttcaatccac 1380
ctgttcttgt ttttgttcag tccattgaaa gggctaaaga actttttcat gagctcatat 1440
atgaaggtat taatgtggat gttattcatg cagagagaac acaacaacag agagataaca 1500
cagtccacag tttcagagca ggaaaaatct gggttctgat ttgtacagcc ttgctagcaa 1560
gagggattga ttttaaaggt gtgaacttgg tgatcaacta tgactttcca actagctcag 1620
tggaatatat ccacaggata ggtcgaactg gaagagcagg gaataaggga aaagcaatta 1680
catttttcac tgaggatgat aagccattat taagaagcgt tgctaatgtt atacagcagg 1740
ctgggtgtcc tgtaccagaa tacataaaag gttttcagaa actactaagc aaacaaaaga 1800
aaaagatgat taagaaacca ttggaaaggg agagcattag tacaactcca aaatgtttct 1860
tagaaaaagc taaggataaa cagagaaagg tcactggtca gaacagcaag aagaaagtag 1920
ctcttgaaga caaaagttaa aaacagactt taaaaatact gtcccagaaa tgtaatttta 1980
tgatcccagc atgaatgtta ttttcatgga atacttgaag tcttacagtc acctgtacca 2040
aacatttgaa atcaactaca agtacatggg actggtgata aatgatccta aactatcaag 2100
tcagtttcaa tttgtaggtg cctttttttt ttcctgtaga gatgagggtc ttgccatgtt 2160
gtccaggctg gtcttgaact cctgacctca cacaatcctc ctgccttagc ctcctgagta 2220
actgagatta caggcacaag ctgctgcacc cagctctgta ggtgactttt aaatgattat 2280
acaatggaaa taacattcat tgacatttct gtggtttgaa tccagggaga tacttcttat 2340
agaaaaacaa atgtttatgc taaaaataac accaaaatgt ggtgaactct taaggacttt 2400
tcccttcaag tgtgaaggaa ggtgtgatga atgctgtgga gaggcatctg gaacagaaat 2460
tcaaaataaa gccttgacat taaatacccc ttccactgct cactttgtgg atggtagcat 2520
gagctgtcta ccaagaagaa acctgctgct ctcttaattt taatatttcc taatttgttg 2580

atggcctttt gtgttgtgaa ccacaacaaa gagaggcctc ttttgtggct ggttattcca 2640
gttccctggg attttaaatt ctttggtcta ttaagtatcc ttgtattgga tacgtaatac 2700
cttagtgctg tcataatgtt gcacaagatc atgatcagct tctccctttc ttcattttct 2760
gtgatttaac catgttcttt cctgtctctt tccatttaag atattttatt tgaatactga 2820
taaacatttt atcctgataa ggaagaatgt tcttgttact tgatatacct ctgtcttcat 2880
tctcttacag cttatctttc cttaggttga tggtgcctca tttaataagt agatctctac 2940
ttaagctaaa ggattatgat catatggtgt tggagactat tagctattag atttttctca 3000
aaatatgagt tttgtacaat ggaatgaaaa agtgacttca tatacgtaag actggtgcct 3060
aagaatgact tgaatgttat cagtactacc acagaactat aatatacatt gcctttctc 3120
agc 3123

<210> 319

<211> 1817

<212> DNA

<213> Homo sapiens

<400> 319
caaccatcct gaagctacag gtgctccctc ctggaatctc caatggattt cagtcgcaga 60
agcttccaca gaagcctgag ctcctccttg caggcccctg tagtcagtac agtgggcatg 120
cagcgcctcg ggacgacacc cagcgtttat gggggtgctg gaggccgggg catccgcatc 180
tccaactcca gacacacggt gaactatggg agcgatctca caggcggcgg ggacctgttt 240
gttggcaatg agaaaatggc catgcagaac ctaaatgacc gtctagcgag ctacctagaa 300
aaggtgcgga ccctggagca gtccaactcc aaacttgaag tgcaaatcaa gcagtggtac 360
gaaaccaacg ccccgagggc tggtcgcgac tacagtgcat attacagaca aattgaagag 420
ctgcgaagtc agattaagga tgctcaactg caaaatgctc ggtgtgtcct gcaaattgat 480
aatgctaaac tggctgctga ggacttcaga ctgaagtatg agactgagag aggaatacgt 540
ctaacagtgg aagctgatct ccaaggcctg aataaggtct ttgatgacct aaccctacat 600
aaaacagatt tggagattca aattgaagaa ctgaataaag acctagctct cctcaaaaag 660
gagcatcagg aggaagtcga tggcctacac aagcatctgg gcaacactgt caatgtggag 720
gttgatgctg ctccaggcct gaaccttggc gtcatcatga atgaaatgag gcagaagtat 780
gaagtcatgg cccagaagaa ccttcaagag gccaaagaac agtttgagag acagactgca 840
gttctgcagc aacaggtcac agtgaatact gaagaattaa aaggaactga ggttcaacta 900
acggagctga gacgcacctc ccagagcctt gagatagaac tccagtccca tctcagcatg 960
aaagagtctt tggagcacac tctagaggag accaaggccc gttacagcag ccagttagcc 1020
aacctccagt cgctgttgag ctctctggag gcccaactga tgcagattcg gagtaacatg 1080
gaacgccaga acaacgaata ccatatcctt cttgacataa agactcgact tgaacaggaa 1140
attgctactt accgccgcct tctggaagga gaagacgtaa aaactacaga atatcagtta 1200
agcaccctgg aagagagaga tataaagaaa accaggaaga ttaagacagt cgtgcaagaa 1260
gtagtggatg gcaaggtcgt gtcatctgaa gtcaaagagg tggaagaaaa tatctaaata 1320
gctaccagaa ggagatgctg ctgaggtttt gaaagaaatt tggctataat cttatctttg 1380
ctccctgcaa gaaatcagcc ataagaaagc actattaata ctctgcagtg attagaaggg 1440
gtggggtggc gggaatccta tttatcagac tctgtaattg aatataaatg ttttactcag 1500
aggagctgca aattgcctgc aaaaatgaaa tccagtgagc actagaatat ttaaaacatc 1560
attactgcca tctttatcat gaagcacatc aattacaagc tgtagaccac ctaatatcaa 1620
tttgtaggta atgttcctga aaattgcaat acatttcaat tatactaaac ctcacaaagt 1680
agaggaatcc atgtaaattg caaataaacc actttctaat tttttcctgt ttctgaaaaa 1740
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800
aaaaaaaaaa aaaaaaa 1817

<210> 320

<211> 1474

<212> DNA

<213> Homo sapiens

<400> 320
ggcagatgaa atataagatt catcaaccac atttgacagc ccatggcagg tttcctgttt 60
tccatcgtcc ctctgcaggt cacagacaca cagagcccag ccgtggcagg ctcagccggg 120
gtccggggct gctaacaacg gctacattcc tcccccaggg ccaagggaaa tcctgagcgc 180

aggccagggt tgtttggttt tgaggtgtgc tgggatgaaa ggcaccctgg aagtggaagg 240
ttcggtcatt cattaattaa ttacatctat aattgagggt ttgttcttaa gagcgagtcc 300
tttgaaagta ctttccttca aacagtgact gccacaaagg catcagatat tcaccacctt 360
ctcggctgcc tcagcacagc aagctttatt ctgggacctg agatcctgtt ctgagctggc 420
tttcccttct ccaggctcgc tcaccctccc tttagagata gtggatggta agatgaccaa 480
tgctcagatt attcttctca ttgacaatgc caggatggca gtggatgact tcaacctcaa 540
gaaatggaga agcatcatgt gccaagtgac ttcaatgtca atgtgaaggt ggatacaggt 600
cccagggaag atctgattaa ggtcctggag gatatgagac aagaatatga gcttataata 660
aagaagaagc atcgagactt ggacacttgg tataaagaac agtctgcagc catgtcccag 720
gaggcagcca gtccagccac tgtgcagagc agacaaggtg acatccacga actgaagcgc 780
acattccagg ccctggagat tgacctgcag acacagtaca gcacgaaatc tgctttggaa 840
aacatgttat ccgagaccca gtctcggtac tcctgcaagc tccaggacat gcaagagatc 900
atctccccact atgaggagga actgacgcag ctacgccatg aactggagcg gcagaacaat 960
gaataccaag tgctgctggg catcaaaacc cacctggaga aggaaatcac cacgtaccga 1020
cggctcctgg agggagagag tgaagggaca cgggaagaat caaagtcgag catgaaagtg 1080
tctgcaactc caaagatcaa ggccataacc caggagacca tcaacggaag attagttctt 1140
tgtcaagtga atgaaatcca aaagcacgca tgagaccaat gaaagtttcc gcctgttgta 1200
aaatctattt tcccccaagg aaagtccttg cacagacacc agtgagtgag ttctaaaaga 1260
taccettgga attatcagac tcagaaactt ttattttttt tttctgtaac agtctcacca 1320
gacttctcat aatgctctta atatattgca cttttctaat caaagtgcga gtttatgagg 1380
gtaaagctct actttcctac tgcagccttc agattctcat cattttgcat ctattttgta 1440
gccaataaaa ctccgcacta gcaaaaaaaa aaaa 1474

<210> 321

<211> 754

<212> DNA

<213> Homo sapiens

<400> 321
caatcaaacc cataaatacc acagactcta atagccatgg attgctgtgc ctctcgaggc 60
tgcagtgtcc ccaccgggcc tgccaccacc atctgctcct ctgacaaatc ctgccgctgt 120
ggagtctgcc tgcccagcac ctgcccacac acagtttggt tactggagcc cacctgctgt 180
gacaactgtc ccccaccctg ccacattcct cagccctgcg tgcccacctg cttcctgctc 240
aactcctgcc agccaactcc aggcctggag accctcaacc tcaccacctt cactcagccc 300
tgctgtgagc cctgcctccc aagaggctgc taatggatgg ctactttgct cagtgcctga 360
gattgaaaaa gtcaacatag aagctttagc attcacctat ctcagtacct acaactaatg 420
tactctgctt tagaaattgg aacaaggatg gtactaccac aatcaccccc tgcaaaaaaa 480
aagagaccaa gaaactttca atgaccattc agctataacc aactgcagtt tgaatcagtg 540
gatgcctata gcttcctgaa gctgttcgat tccttcatat taaagtgtct ctttctgtgg 600
gtggtttggg aattctgttt tcagtcttgg gtggtatctt tctgaaaatt aaggaagttc 660
ttcatgatta tcctaataaa ttttacatct ctggcatagc acaaaaaaaa aaaaaaaaaa 720
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 754

<210> 322

<211> 749

<212> DNA

<213> Homo sapiens

<400> 322
aagaaactga aagctaacca gacgcccatt gccatggatt gctgtgcctc tcgcagctgc 60
agtgtcccca ctgggcctgc caccaccatc tgctcctccg acaaatcctg ccgctgtgga 120
gtctgcctgc ccagcacctg cccacacaca gtttggttac tggagcccat ctgctgtgac 180
aactgtcccc caccctgcca cattcctcag ccctgcgtgc ccacctgctt cctgctcaac 240
tcctgccagc caactccggg cctggagacc ctcaacctca ccaccttcac tcagccctgc 300
tgtgagccct gcctcccaag aggctgctga tggatggcta ctttgctcag tgcccgacaa 360
cgaagaatcc agaagctgtc ccttcagtat tcacttgcct cagtagtttg ccagatgtta 420
aggtagacca gatgacccag atatgaagaa cttacctttg gttttaatgg gggaaaaaaa 480
gaaaagtatt ttttatggtt atttagctga aaaaccattt ggttcctgtg ggcaggtgaa 540

tgagttttat tagcaaaata ctgtttcaat ctttaagacc tcagattaca tgttcttgat 600
catattgctt cctggctctt gtttcttgta ctgggtattt tcatagaaga aaatttcttg 660
gtgggttttc caataaacta tatttctctg gcaaaaaaaa aaaaaaaaaa aaaaaaaaaa 720
aaaaaaaaaa aaaaaaaaaa aaaaaaaaa 749

<210> 323

<211> 440

<212> DNA

<213> Homo sapiens

<400> 323
gtgttccctg agataggtgg atataaaaga cccatagagg acaaccttgt agaagaaagc 60
cttcctttgc cacaaaacat tgtcctggtc cactgtcact atgtcttgct gtgattccta 120
tctccaagga tgctgcagcg tccccactgg cctggccacc actatctgcc cctctgacat 180
aagctgtcaa tgtgaagtct gcctacccag cacctgtcct catgagatca gcctccttca 240
gcccacctgc tgtgaacctg gcccctgcct ggctgcatgc ctgactccta tgtgccatcc 300
tgttgactgc tcaacaaatg ccacccagct ccaaccctga gcgggctctc tgtcaccacc 360
tgcatccaga gtgtgaacca ccttgctgct agccaaagag cttgcccaca ttaccctgag 420
gaccttcagt agtcattaag 440

<210> 324

<211> 614

<212> DNA

<213> Homo sapiens

<400> 324
agacttctct caactcaaca aaaacccacc tcccattgcc atgtattgct gtgctctccg 60
ctcctgcagc gtccccaccg gccctgccac caccttctgc tcatttgata aaagctgccg 120
ctgtggagtc tgcctaccca gcacctgccc acatgagatc agcctccttc agcccatctg 180
ctgtgacacc tgcccccccac cctgctgcaa gcctgatacc tatgtgccaa cttgctggct 240
gctcaacaac tgtcacccga ctcccggact gagtgggatc aacctgacca cctatgttca 300
gcctggctgt gagagtccct gtgagccccg ctgttaacca gccgagtctg cacaggttcc 360
gtgaggtggc tgcccaatgt cctctgcacc atctgggctt cagcactcac tactgcctac 420
atcaaggcta aggccatccc aatccccggg gccaagtctt gatgaatctt cttaattatt 480
tgcacatttg ggtaccattg gagacctccc ttctgtcttt taggctattt catcactctt 540
tgagaaataa ccatttttgac catttgttaa taaactttat tctggcttag caaaaaaaaa 600
aaaaaaaaaa aaaa 614

<210> 325

<211> 1193

<212> DNA

<213> Homo sapiens

<400> 325
cagaaactcc tccaagcaac ctaactctta acccaacttc tgacaccatg acctgctgcc 60
agaccagctt ctgtggatat cccagcttct ccatcagtgg gacctgtggc tccagctgct 120
gccagccaag ctgctgtgag accagctgct gccagccacg cagctgccag actagcttct 180
gcggatttcc cagcttctca accagtggga cctgcagctc cagttgctgc cagccaagct 240
gctgtgagac cagctgctgc cagccaagct gctgtgagac cagctgctgc cagccaagct 300
gctgccagat cagctcctgc ggaactggct gtggcattgg tggtggcatc agctatggcc 360
aggagggcag cagtggagct gtgagcaccc gtatcaggtg gtgccgccca gacagtcgtg 420
tggagggcac ctacctaccc ccctgctgtg tggtgagctg cacgcccccca tcctgctgcc 480
aactgcacca tgcccaggcc tcctgctgcc gcccgtccta ctgtggacag tcctgctgcc 540
gcccagtctg ctgctgtgag cccacttgct gaaagccagt ttgcttattt tcaattgcct 600

aggtcacagt gtctctgaac tgttcatccc ttgaccacct ctggaccact aacaagttct 660
cagactttgc attgcttgtg atggagacta ctaagtatat gagctcacaa ttctatctga 720
ttccattcta caatgaatac cttgaccctt cactggggac acagaaatgc tacaaagcca 780
cctgctgatc atcaatttgc ttgggatata ctatttctga tatttctgca ggattaaaaa 840
ttactgacat gttgtggaat ttatccatga gaactatcca caagtctaat gtttccatgc 900
tttataatct attttatctt gtttacctaa aattttttgc aacatcaaag acaccaaatt 960
atagccaagt gacattcctc aagtcaccag agagaatgga agctcatcac ccaacattca 1020
gcttctaaga agtaggctgg actttccaca ttttaacatc tgatccatcc cttggttttt 1080
ggatcataat gatcttgcct gctggatatt tcagttatat ctgtgataca atgtcttctg 1140
tcatttctta ataaatatta tatactaggc aaagaaaaaa aaaaaaaaaa aaa 1193

<210> 326

<211> 986

<212> DNA

<213> Homo sapiens

<400> 326
aagcaaccca gacttcatac cagctcccaa caccatgacc tgctgccaga ccagcttctg 60
tggatatccc agctgctcca ccagtgggac atgcggctcc agctgctgcc agccaagctg 120
ctgtgagacc agctgctgcc agccaagctg ctgccagacc agcttctgcg gatttcctag 180
cttctcaact agtgggacct gcagctccag ttgctgccag ccaagctgct gtgagaccag 240
ctgctgccag ccaagctgct gccagaccag ctcctgcgga actggctgtg gcattggtgg 300
tggcattggc tatggccagg agggcagcag tggagctgtg agcacccgta tcaggtggtg 360
ccgcccagac tgccgtgtgg agggtacctg cctgcccccc tgctgtgtgg tgagctgcac 420
acccccaacc tgctgccagc tgcaccacgc cgaggcctcc tgctgccgcc catcctactg 480
tggacagtcc tgctgccgcc cagtctgctg ctgctactcc tgtgagccca cctgctaaaa 540
gccagtttgc tgattttcaa cttgaaattt ccactttcag ttccattcat gaacgaatta 600
tttcttcaag cacttatgga caacgaacaa attcttcaac ctttctttgt ctttcttatg 660
gggttacca aatattttgg cctcagaatt atctgattcc tttcaattcc agaaagacct 720
tactcttctc tctgaggacg ccaaaataca aatttgaccc aagaaatgaa aaagccgatt 780
taccttgaaa ctgagccttt gcaagcattg aagcccacgc tctgagtctc agcgccgacg 840
agaccatgga agagccatct gtccttctca ggacactcac ttcctgtatc ccaccgtcct 900
gcaaattgca ccccctatga aagaggaata atataccaag gtctaataaa ttttaactat 960
tggtgcaaca aaaaaaaaaa aaaaaa 986

<210> 327

<211> 903

<212> DNA

<213> Homo sapiens

<400> 327
aataggcagc cataattcag aaactcctcc aagcaaccca accttcagat caactcctga 60
caccatggcc tgctgtcaga ccagcttctg tggatttccc agctgctcca ccagtgggac 120
ctgcggctcc agctgctgcc agccaagctg ctgtgagacc agctcctgcc agccacgctg 180
ctgtgagacc agctgctgcc agccaagctg ctgccagacc agcttctgtg gatttcctag 240
cttctcaacc ggtgggactt gtgactctag ctgctgccag ccaagctgct gtgaaactag 300
ctgctgccag ccaagctgct accagaccag ctcctgcgga actggctgtg gcattggtgg 360
tggcattggc tatggccagg agggcagcag tggagctgtg agcacccgta tcaggtggtg 420
ccgcccagac tgccgtgtgg agggtacctg cctgcccccc tgctgcgtgg tgagctgcac 480
acccccatcc tgctgccagc tgcaccacgc cgaggcctcc tgctgccgcc catcctactg 540
tggacagtcc tgttgccgcc cagtctgctg ctgctactgc tctgagccca cttgttgaaa 600
acctccttct gctggggatc ctgataagat ggcaccttaa aactagccaa attagaatcc 660
taacaatctt ctgaactcca gtacctataa ctgggcttgc aacctctcat cacacagcca 720
cataaattcc ctaggaagta aattcattta caatggaaga ccaaaaattt ttcctagacc 780
tggttgtcag ccaaagtcct acaatgtgaa aagagttaga tactatttta ctataaatat 840
cacctgaaat atttcaacag ttattgggac ttaaatttaa taaaagtttt catctcttca 900
atg 903

<210> 328

<211> 615

<212> DNA

<213> Homo sapiens

<400> 328
cgtgacctgc gtgccccgct gcacgcgccc catctgcgag ccctgccgcc gcccggtgtg 60
ctgcgacccc tgctccctgc aggaaggctg ctgccgcccc atcacctgct gcccctcgtc 120
gtgcacggct gtggtgtgca ggccctgctg ctgggccacc acctgctgcc agcctgtgtc 180
tgtgcagtcc ccctgctgcc ggcccccctg cggccagccg acccccttgca gcaccacctg 240
caggacctcc tcctgctgag ccaccacctg ctgccagcct gtgtctgtgc agtcccccctg 300
cggccagccg acccccttgca gcaccacctg caggacctcc tcctgctgag cagcccgtta 360
tcacgaaggg cccctcagaa gatggccagg tccatcccgc tgcccctcag ggcttcaccg 420
cagagcaata cacgtttcct tgagaagccc atttctcatc tcttcatact agctcacact 480
atgcattgaa gacacctttt cagaccaacc acagatgaga aatacttttc ctaggactcc 540
agtctaactc ctatatcatg ttgtctgctt tctaataaac tcaatactcc taccataaaa 600
aaaaaaaaaa aaaaa 615
<210> 329

<211> 812

<212> DNA

<213> Homo sapiens

<400> 329
aatagcccaa cccacaccag cctcagacac caccatgacc ggctcctgct gcggctccac 60
cttgtcctcc ctgagctacg gggaggctg ctgccagccc tgctgctgcc gcgacccctg 120
ctgctgccgc cccgtgacct gccagaccac cgtgtgccgc cccgtgacct gcgtgccccg 180
ctgcacgcgc cccatctgcg agccctgccg ccgcccggtg tgctgcgacc cctgctccct 240
gcaggaaggc tgctgccgcc ccatcacctg ctgcccctcg tcgtgcacgg ctgtggtgtg 300
caggccctgc tgctgggcca ccacctgctg ccagcctgtg tctgtgcagt cccccctgctg 360
ccggcctccc tgcggccagc cgacccccttg cagcaccacc tgcaggacct cctcctgctg 420
agaccccaat gcccccacag agcaatacac tgaagcctaa acatctatct ggtgttttta 480
aaaagttaaa agaaaaatag atttttttttc acaaggtgac aatagtgatt tttaccatct 540
ggatacagcc tggtgtaagc agacgtccat taccaccctc acccacattt tcaggtgtct 600
acatcagcct tagtcattat ggatagtaaa tcgaccttta agaattcctg gggtggactt 660
tgcaaacaca ttctacaacc tgatggtttt tactgctcaa actgtcacca tcatcttttg 720
caatgtgttg ctcactgttg tcaataaact aatttttcct ggcaaaaaaa aaaaaaaaaa 780
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 812
<210> 330

<211> 601

<212> DNA

<213> Homo sapiens

<400> 330
tgcttggtga gtgtggctac ttccttgaac cacatggaga cacattgtac aggaagaagt 60
gcctccttct gctcgtcatc tgcgattctc atctgacacc atggtcagct tctgttgtag 120
ctctgtctgc tctgaataga gctgtggcca aggcctctgc cagacctgct gctgctgcag 180
ttgctgccag accacctact gcagaaccat ctgctaccat cccagctgct ctgtgtccag 240
ctgttgcagg cccccgtgct gccagtccct gtgctgcccc agctgctgca tttctagccg 300
ctgccactca agctgcagtg tgtccacctg ctccaggccc agctgttata atccccagta 360
ctaccagccc tcctgctgcc acccttctgc tgcacttcta gctgctgcca ccctggctgc 420
tgtgtgtcca gctgctgctg tccagtctgc taccagacca cctgctgtgg tccagtctcc 480
tatgaatcct cttgctgttg aacttcattc ctgaccacca gccctggttc aaccaccttg 540

ttgtcagtgt accagtcatt ctcattcccc ttctccactg gacctggcct tgccctgatc 600
c 601

<210> 331

<211> 1202

<212> DNA

<213> Homo sapiens

<400> 331
cttcactctc ttgaaaaccc acccagatcc tccccgttct gacaccatgg tcagctcctg 60
ttgtggctcc gtgtgctctg accagggctg cggccaagtc ctctgtcagg agacctgctg 120
ccgccccagc tgctgtcaga ccacctgttg caggaccacc tgctaccgcc ccagctgttg 180
tgtgtccagc tgctgcaggc cccagtgctg ccagtctgtg tgctgccaac ccacctgctg 240
tcgccccagc tgctgtgaga cgacctgctg ccaccctagg tgctgcatct ccagctgctg 300
ccgccccagc tgctgtatgt ccagctgctg caagccccag tgctgccagt ctgtgtgctg 360
ccagcccacc tgctgccgcc ccagctgctg catctccagc tgctgtcgcc ccagctgctg 420
tgtgtccagg tgctgcaggc cccagtgctg ccagtctgtg tgctgccagc caacctgctg 480
ccgtcccagc tgctgcatct ccagctgctg ccgcccctct tgctgtgaat ccagctgctg 540
ccgcccatgc tgctgccgcc cctgctgctg cctgcgtcca gtctgtggcc gagtctcctg 600
ccacaccact tgctatcgcc caacctgtgt catctccacc tgtccccgcc ccttgtgctg 660
tgcctcctct tgctgctgag cccactgccc tggctcacgt ccccttcac cactggccca 720
cagatgtaga cccttctact gtgctgacca ttaggataca tgaagtgggg ttgatgtcat 780
tcaataggat ggaccttatg cttccaaaga gcccaccacc atttcactga ctctgtgaga 840
acattctggt tcattttaaa ctccctccct tgctttcttt ttcttctggt ggtggcacca 900
aatgtgaatt aatttgtaat acactagcta agaaattatt ccaatcttct gatttcctta 960
ttttctttat cactttaagg tacagattct ccttctcagt gaggtagata ttatctgcag 1020
gaccagtttt gtcactgatg ttgcaccctc agatccagcc acccaattgt attctgtgtt 1080
tctcctaggg tgaatttctt atgctttgtt gcatctctgc tttctaataa acttttctgc 1140
acttaagaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1200
aa 1202

<210> 332

<211> 1219

<212> DNA

<213> Homo sapiens

<400> 332
ttcactctcc tagaaaccca actagatcct tcaccctctg acaccatggt aaactcctgt 60
tgtggctccg tgtgctctca ccaaggctgt ggccaagacc tctgccagga gacctgctgc 120
cgccccagct gctgtgagac cacctgctgc aggaccacct actgtcgccc cagctgctgt 180
gtgtccagct gctgcaggcc ccagtgctgc cagtctgtgt gctgccagcc cacctgctgc 240
cgccccagat gctgcatctc cagctgctgt cgccccagct gctgtgtgtc cagctgctgc 300
aagccccagt gctgccagtc tatgtgctgc cagcccactt gctgccgccc cagatgctgc 360
atctccagct gctgtcgccc cagctgctgt gtgtccagct gctgcagacc ccagtgctgc 420
cagtctgtgt gctgccagcc cacctgctgc caccccagct gcagcatctc cagctgctgc 480
cgcccctctt gctgtgaatc cagctgctgc cgcccctgct gctgcctgcg tccagtctgt 540
ggcggagtct cctgccacac cacttgctat cgcccaacct gtgtcatctc cagctgcccc 600
cgccccttgt gctgtgcctc ctcttgctgc tgagcccact gccctggctc atctctccct 660
tcactgcagg cccacagttg tagaccattc ttctgtgctg actattagga cacatggagt 720
gggattgatg tcattcagca gggtggactt catgtttcca atgagcccat caccatccca 780
ctgactctgt gagaacattc tggttcattt taaactccct ccccttgcttt ctttctcttc 840
cagtcatggc accaaatatg aaataatttg taatccacta gctaagaaat tattccaatc 900
ctctaaattc ctcatttttt aaatcgtttt gagcctacag aatatccttc ccagtgaggt 960
acacattatc tccatttcaa acatattatt tgtctgtcag cctttcagtc attcttttct 1020
cttggaaagg taggaggctg cccctcccat gctctcctgc tttctccctg ttctctcttt 1080
gtctctgttt gttcaagttt gccagaattt ttctatttta ttggttcttt atctttattg 1140
tgttcacaaa atatattgta ttaaactttt catttagaaa aaaaaaaaaa aaaaaaaaaa 1200
aaaaaaaaaa aaaaaaaaa 1219

cttccctagg gtcctcaccc tctctgtttt ccaaaacctc cccccattca cattacatat 2460
gtctggcccc cagagacata tgacacagtg tccaacataa gaatttaggc cacctaggac 2520
catgccccac actccccaat gcctcttcag gccccaggcg tcctgcttgg taagagccta 2580
gcctctatga cctccccagc taggaagccc ttcccacctt gaccatctgt tcaatcaaat 2640
gttcaccaag ccaagggctg gctatgagga catcaataaa taagccatag gccctgccct 2700
gagggagctt atagtgagag gcattctccc tactctatcc ctcccaacac tgctctacat 2760
tcctttcctc caggaagtct tccctaattg tcactacccc ctacccccat ctgtccgccg 2820
ctagcctgac agtaacccag cacagcccac aggcctctca gtctgtcgct catcccatcc 2880
aggcccacct ggacagaccc tcacttagcc ctgaataagc aattatgaaa caaggattct 2940
ccagctgcat cctcaggcca ggagagacca gcaggggcca tgtggtctac cccttgcctc 3000
tgcacatctt actgccagtt ctgtcttgac tctggctgtt ccccaaggag ggctggggat 3060
attggctcct ccaaccaaca tttctcccac ttccacacca ggtgcacttc ctggagcagc 3120
agaacaagct gctggagaca aagctgcagt tcttccagaa ccgtgagtgc tgcgagagca 3180
acctggagcc tctgtttgag ggctacaact ctgcgacggg aggccgagtg catggaagcc 3240
gacagtggga ggctggcctc agagctcaac cacgtgcagg aggtgctgga gggttacaag 3300
aagaagtaag tgcagtggga ggagggtttg gagaaattaa atgggattaa gtcaaagtat 3360
gatagatttg ggttcgacca taagaggaac ttcccaacag gaagaaaata gtacaataat 3420
cctgtggggt gggggaattt tcttttctag actgggagtt taaacattgg agattctgaa 3480
ctgctccagc tacttcatca ttatgtttat ttcattaatt catcaaacat tttattgagc 3540
atctactcaa tgttgtaggt gctatgttgg gctctgagtg ggaaaggagt aaggaagtaa 3600
gaaggaggag aaaggagagg taccagaact gggcagggtt cctgccccaa aggagttaca 3660
gcccagaaac agagaggagg tcttcgaaaa tgtgcagggc ccccaccctg acccacatca 3720
ggaatgaaag tccctgaaat agaaagagct ttgagtaaca tgggttggtt tgaatgccca 3780
gaagaaaggg atggacaagt tcccttcagt ccaacacttt tttgattctt gtgaccacct 3840
ccaaaagatc cagttcccgg aatgtgtttg cctgtgcctt cattcccatt ttttttcctc 3900
cccagatctg agccttgtgt gttttgtccc caggtatgaa gaagaggtgg cactaagggc 3960
cactgctgag aatgaatttg tggccttgaa gaaggtgagt gaccccagaa tccttgtccc 4020
ctaactcaga atggaggcca gagacagaag ggtcctagac ccaaaggttt gagctgggag 4080
accttgacta acatcaggct ctctgtagtt cggagaggtg gagggactcc cttagaatca 4140
cacggcccag cttctttctg ttcctttccc cagtgctcca gtgctccctg ggtgcctctg 4200
gcagcagaat acactgcctg cttcaggcag ggcaatattc tcccccatct gaggttgctg 4260
ccatatagcc accagatttt cacctccttc agccctgcgc ctggcataga gtaggcaccc 4320
tgtaaatatt tgcgagtgaa agagcaggtg aattatagac aatggctgca tcctcaggcc 4380
agcaacagct cagaaccaca ggcaggggca gtgaccgtca ttccttgcct cttagacttg 4440
actctgtaca agtgcctgga ctctcattag gcccttccag ccacctccat atgaggaaag 4500
caggggttat tgtgagaaaa tgaggtgcag aggttaaggg ggctgctcag agtgacactg 4560
acttggaagg gacagagttg agccctgagc cctcttctga ccacagatct agattatctt 4620
ccatgagatt ccagagcgga gaggggtggg gcatcattag aatttaccca ctgtacacca 4680
gatcttaagt tttggtgata gtgcatgaca tggattctag agaaatcatc ccagtgggtg 4740
ccagaggaag tacctgcaaa cctcagtgtc cccctccctt acccaaacct gctcttgacg 4800
ccactactaa ctgtataagg ggcttcacca cggcatgata ccaggcactg ggatgagcac 4860
gtacagaggg tttccaggga agctgctcag cctgatgggg tgcccctccc tggtaggaaa 4920
aaaacaaggg ttctgaattg tctcccagtg caattcctgt gccaccttct cccctttagg 4980
atgtggactg tgcctacctg cgcaaatcag acctagaggc caacgtggag gccctgaccc 5040
aggagatcga cttcctgagg cggctatatg aggaggtgcg gggtcacaaa gcagcggggc 5100
agcaaggagg atatgagtta gacattggaa aatacttgtc accattcagc ttctgaccct 5160
atggtgggac aaggggagaa accccaggca ggttcaccat gagcaggtga agaagtgtgg 5220
cttgacctcc cactgccatg ccccaggcaa cctggtgttc ctggttacta aagtgacaat 5280
tgccacaagg tggtaaactc tggtcctaga cccaaaggtt caagctggaa ggtcctgatt 5340
aaaatcaggc actgggcctc agacatggac tctctacctc caagtcctct aggctcagga 5400
ctctgggctt gaggagtcct gagggggtcc cagggttgag gtggaagact gtaactggga 5460
gggtggggct tcagacatcg aagaagcagg gaaaggggac aagggatgat gatgagggtg 5520
ggctaccaag agtagggctc tggtctaaga agctcttgcc actccccatc tcctgcagga 5580
gatccgcgtt ctccagtcgc acatctcaga cacctcagtg gtcgtcaaga tggacaacag 5640
tcgggacctg aacatgcact gtgtcatcac tgagatcaaa gctcagtacg atgacattgc 5700
cacccgcagc cgggccgagg ccgagtcctg gtaccgcagc aaggtgagtg gcacaggaca 5760
cctgcctgct agacatggca gttggaggga tgcaaggtac ctattaaata ggcttccttt 5820
tctggggatt ctggtcccta cggatgggaa gataagggct gcctctctga ggttggggta 5880
gcagggcagg actgccatgt gtggttgcac aggctgagca ctgcacaacc tgcacaatca 5940
tccgtggtgt cctgaatgga tgggaggtcc caccctgagc ctcataagca actctacttc 6000
cccagtgtga ggagatgaag gccacggtga tcaggcacgg ggagaccctg cgccgcacca 6060
aggaggagat caacgagctg aaccgcatga tccagaggct gacagccgag gtggagaacg 6120
ccaagtgcca ggtatgggag cccagcccac ccctcagtgt aggaagagag tggggttagc 6180
cctcagcaga gggatttggg ttagacccct tgccagtgtg tggggacaag ggcaaacacc 6240
cagcatgagg acagacagcc tcctccatgg aaaacgttct ccattctcgc tgaatggagg 6300
cattggggca atatggaggg aatgaggtgc ttcaaggcca gagtggggca gaggaaatag 6360

```
cggcatcatc gctgaggtca aggcacagta tgaggagatg gccaaatgca gccgggctga    960
ggctgaagcc tggtaccaga ccaagtttga gaccctccag gcccaggctg ggaagcatgg   1020
ggacgacctc cggaataccc ggaatgagat ttcagagatg aaccgggcca tccagaggct   1080
gcaggctgag atcgacaaca tcaagaacca gcgtgccaag ttggaggccg ccattgccga   1140
ggctgaggag cgtggggagc tggcgctcaa ggatgctcgt gccaagcagg aggagctgga   1200
agccgccctg cagcgggcca agcaggatat ggcacggcag ctgcgtgagt accaggaact   1260
catgagcgtg aagctggccc tggacatcga gatcgccacc taccgcaagc tgctggaggg   1320
cgaggagagc cggttggctg gagatggagt gggagccgtg aatatctctg tgatgaattc   1380
cactggtggc agtagcagtg gcggtggcat tgggctgacc ctcgggggaa ccatgggcag   1440
caatgccctg agcttctcca gcagtgcggg tcctgggctc ctgaaggctt attccatccg   1500
gaccgcatcc gccagtcgca ggagtgcccg cgactgagcc gcctcccacc actccactcc   1560
tccagccacc acccacaatc acaagaagat tcccacccct gcctcccatg cctggtccca   1620
agacagtgag acagtctgga aagtgatgtc agaatagctt ccaataaagc agcctcattc   1680
tgaggcctga gtgatccacg tgaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa   1740
aaaaaaaaaa aaa                                                      1753
```

<210> 365

<211> 9721

<212> DNA

<213> Homo sapiens

<400> 365

```
aaactccctt cccctcgccc aagcagagtg gagaggggta caggaaggag ggacgcttac     60
gagaagagga ggtatgagct gagtcctgag atagaagaga agagaaagtt tttgcccaaa    120
gggacacagg atgaatgata aattaacaat agaataaata gaacaaataa tgaataaata    180
aacatttaca cagttcttat taagtgatag gtagggccct aaatgctttc tggtcattat    240
ctcattagat cctggcaacg gctctcagtg cttactatta ttatcgctat tttgcctgtg    300
agaaaactga ggctcagaaa ggttattcag ctgcccaagg tcatgacatg aacctgcggg    360
actgacttca gaggccccac actctcgaca cacgcacaca ctcgccccac acacggtgcc    420
cgctgcgggg gaccggggct atgggactgg gaccgaagtg gaggccggat ccaacaaggc    480
aaagctcgga gtcccggcgg aggggcggac aggtggaggg tttcactcgc actgaccgca    540
gcgcacgggc agcggggctg accaggcgct tcccggcccc ataaaactcg ggcgggaggc    600
gaacgccgtt ttatgcaggc ggccgggcgg cctggctcgc tttcattccc ggctccgcga    660
ggcgcgaggg ccgagccctc caagctcgta aacgccttgg ccgcgagctc cctccccggg    720
cccgcgggcc gtcatataag gcgcagcgcc ggcttggggt ccagccgccc gcccctgccg    780
ccaccgcacc atgtcctgcc tctactcccg cctcagcgcc ccctgcgggg tccgcgcctt    840
cagctgcatc tcggcctgcg ggccccggcc cggccgctgc tgcatcaccg ccgcccccta    900
ccgcggcgtc tcctgctacc gcggcctcac cggaggcttc ggcagccaca gcgtgtgcgg    960
gggcttccgc gccggctcct gcggacgcag cttcggctac cgctccgggg gcgtgtgcgg   1020
acccagcccc ccatgcatcg ccaccgtgtc ggtcaacgag agcctcctca cgcccctcaa   1080
cctggagatc gaccccaacg cacagtgcgt gaagcaggag gagaaggagc agatcaagtc   1140
cctcaacagc aggttcgcgg ccttcatgga caaggtgggt gtcctggatc acacctttcc   1200
tgaacctgga agtcctggaa ggagatggaa attcccagtg gtctctctag aaatggggaa   1260
tgaggatggc tgatatttta tataatgtta ggcggtgccg gcccataata ttaaataaca   1320
ttgaaacacg tggtaagaaa atacatcttc tttaatgtct gtttcaatcc ttaagattag   1380
gtgggaaaaa aatatcactg gatgctaatg tttttagcat ctctctaata cttgcttccc   1440
tcctgcttct aataaagaga gcaagccccg gcttcagaac tgtctaggac aactgtttct   1500
actagaattt aatgttctgt aatttccctt tatgtatgcc tagtaatgat ggttttgcat   1560
ctctggtggc catatgaagt tactctttta accccatacg tgtaaatgtt aattgaatcc   1620
acttatagat aaattaggag taaaatagta tatgggatgg caaaatcctg aagtcagcct   1680
tgggtgtaca cagaagagaa aggaggcaga gagagggtga ggggctggcc caagggccca   1740
cagagagtgg gtgccagaac agggttccaa gcccaggccc tctccaactc tccagcccct   1800
gggctagaag gtccacagct tgggatccca gctgccccac tccctctaga aggggctggt   1860
tcacctcaac aggcactgca cgctggagag gtgacccagg agcggctcgt ggaatggaca   1920
ctggcccagg agcaggccac ctggcctcca tgctcagtta agttcctcta ctgacgcatt   1980
gtgtggcact gggcagaggc ttcccttctc caccccagtg gcttcctcac cagggctgtg   2040
tgtggtgcag tgtgtgagct ggccatgact ctgaaaggga aaatggggcc caaggcagcc   2100
atgagacccc atagccaaga ccacctgaaa ttatgccatg taatatcagc agtccccact   2160
ccccatgtct gcaagttcca acctgggaat tcccagctgc actcatggag ttggatgaaa   2220
aggaccccag ggcttaaagg gaaggagaca tctcagctcc agctcaaatg tacgagaagt   2280
ggaaaactgg aagccaaggg ccctagtttg tttgctgaca tcctttctgg ccagcacaat   2340
gtattgttgc catgtgttcc tccagtttca gtgaaatcca ggaaaagcgt acaagcctca   2400
```

cgggtcagga ccctgctctc ctggcgcagt tcccagctat ctcccctgct cctctgctgg 1620
tggtgggcta ataaagctga ctttctggtt gatgcaaaaa aaaaaaaaaa aaaaaaaaaa 1680
aaaaaaaa 1688

<210> 363

<211> 1512

<212> DNA

<213> Homo sapiens

<400> 363
ctcctctcca gcccttctcc tgtgtgcctg cctcctgccg ccgccaccat gaccacctcc 60
atccgccagt tcacctcctc cagctccatc aagggctcct ccggcctggg gggcggctcg 120
tcccgcacct cctgccggct gtctggcggc ctgggtgccg gctcctgcag gctgggatct 180
gctggcggcc tgggcagcac cctcgggggt agcagctact ccagctgcta cagctttggc 240
tctggtggtg gctatggcag cagctttggg ggtgttgatg ggctgctggc tggaggtgag 300
aaggccacca tgcagaacct caatgaccgc ctggcctcct acctggacaa ggtgcgtgcc 360
ctggaggagg ccaacactga gctggaggtg aagatccgtg actggtacca gaggcaggcc 420
ccggggcccg cccgtgacta cagccagtac tacaggacaa ttgaggagct gcagaacaag 480
atcctcacag ccaccgtgga caatgccaac atcctgctac agattgacaa tgcccgtctg 540
gctgctgatg acttccgcac caagtttgag acagagcagg ccctgcgcct gagtgtggag 600
gccgacatca atggcctgcg cagggtgctg gatgagctga ccctggccag agccgacctg 660
gagatgcaga ttgagaacct caaggaggag ctggcctacc tgaagaagaa ccacgaggag 720
gagatgaacg ccctgcgagg ccaggtgggt ggtgagatca atgtggagat ggacgctgcc 780
ccaggcgtgg acctgagccg catcctcaac gagatgcgtg accagtatga gaagatggca 840
gagaagaacc gcaaggatgc cgaggattgg ttcttcagca agacagagga actgaaccgc 900
gaggtggcca ccaacagtga gctggtgcag agtggcaaga gtgagatctc ggagctccgg 960
cgcaccatgc aggccttgga gatagagctg cagtcccagc tcagcatgaa agcatccctg 1020
gagggcaacc tggcggagac agagaaccgc tactgcgtgc agctgtccca gatccagggg 1080
ctgattggca gcgtggagga gcagctggcc cagcttcgct gcgagatgga gcagcagaac 1140
caggaataca aaatcctgct ggatgtgaag acgcggctgg agcaggagat tgccacctac 1200
cgccgcctgc tggagggaga ggatgcccac ctgactcagt acaagaaaga accggtgacc 1260
acccgtcagg tgcgtaccat tgtggaagag gtccaggatg gcaaggtcat ctcctcccgc 1320
gagcaggtcc accagaccac ccgctgagga ctcagctacc ccggccggcc acccaggagg 1380
cagggagcag ccgcccccatc tgccccacag tctccggcct ctccagcctc agccccctgc 1440
ttcagtccct tccccatgct tccttgcctg atgacaataa agcttgttga ctcagctaaa 1500
aaaaaaaaaa aa 1512

<210> 364

<211> 1753

<212> DNA

<213> Homo sapiens

<400> 364
cagccccgcc cctacctgtg gaagcccagc cgcccgctcc cgcggataaa aggtgcggag 60
tgtccccgag gtcagcgagt gcgcgctcct cctcgcccgc cgctaggtcc atcccggccc 120
agccaccatg tccatccact tcagctcccc ggtattcacc tcgcgctcag ccgccttctc 180
gggccgcggc gcccaggtgc gcctgagctc cgctcgcccc ggcggccttg gcagcagcag 240
cctctacggc ctcggcgcct cgcggccgcg cgtggccgtg cgctctgcct atgggggccc 300
ggtgggcgcc ggcatccgcg aggtcaccat taaccagagc ctgctggccc cgctgcggct 360
ggacgccgac ccctccctcc agcgggtgcg ccaggaggag agcgagcaga tcaagaccct 420
caacaacaag tttgcctcct tcatcgacaa ggtgcggttt ctggagcagc agaacaagct 480
gctggagacc aagtggacgc tgctgcagga gcagaagtcg gccaagagca gccgcctccc 540
agacatcttt gaggcccaga ttgctggcct tcggggtcag cttgaggcac tgcaggtgga 600
tgggggccgc ctggaggcgg agctgcggag catgcaggat gtggtggagg acttcaagaa 660
taagtacgaa gatgaaatta accgccgcac agctgctgag aatgagtttg tggtgctgaa 720
gaaggatgtg gatgctgcct acatgagcaa ggtggagctg gaggccaagg tggatgccct 780
gaatgatgag atcaacttcc tcaggaccct caatgagacg gagttgacag agctgcagtc 840
ccagatctcc gacacatctg tggtgctgtc catggacaac agtcgctccc tggacctgga 900

<400> 361
acccgagcac cttctcttca ctcagccaac tgctcgctcg ctcacctccc tcctctgcac 60
catgactacc tgcagccgcc agttcacctc ctccagctcc atgaagggct cctgcggcat 120
cgggggcggc atcgggggcg gctccagccg catctcctcc gtcctggccg gagggtcctg 180
ccgcgccccc agcacctacg gggcggcct gtctgtctca tcctcccgct tctcctctgg 240
gggagcctac gggctggggg gcggctatgg cggtggcttc agcagcagca gcagcagctt 300
tggtagtggc tttgggggag gatatggtgg tggccttggt gctggcttgg gtggtggctt 360
tggtggtggc tttgctggtg gtgatgggct tctggtgggc agtgagaagg tgaccatgca 420
gaacctcaat gaccgcctgg cctcctacct ggacaaggtg cgtgctctgg aggaggccaa 480
cgccgacctg gaagtgaaga tccgtgactg gtaccagagg cagcggcctg ctgagatcaa 540
agactacagt ccctacttca agaccattga ggacctgagg aacaagattc tcacagccac 600
agtggacaat gccaatgtcc ttctgcagat tgacaatgcc cgtctggccg cggatgactt 660
ccgcaccaag tatgagacag agttgaacct gcgcatgagt gtggaagccg acatcaatgg 720
cctgcgcagg gtgctggacg aactgaccct ggccagagct gacctggaga tgcagattga 780
gagcctgaag gaggagctgg cctacctgaa gaagaaccac gaggaggaga tgaatgccct 840
gagaggccag gtgggtggag atgtcaatgt ggagatggac gctgcacctg gcgtggacct 900
gagccgcatt ctgaacgaga tgcgtgacca gtatgagaag atggcagaga agaaccgcaa 960
ggatgccgag gaatggttct tcaccaagac agaggagctg aaccgcgagg tggccaccaa 1020
cagcgagctg gtgcagagcg gcaagagcga gatctcggag ctccggcgca ccatgcagaa 1080
cctggagatt gagctgcagt cccagctcag catgaaagca tccctggaga acagcctgga 1140
ggagaccaaa ggtcgctact gcatgcagct ggcccagatc caggagatga ttggcagcgt 1200
ggaggagcag ctggcccagc tccgctgcga gatggagcag cagaaccagg agtacaagat 1260
cctgctggac gtgaagacgc ggctggagca ggagatcgcc acctaccgcc gcctgctgga 1320
gggcgaggac gcccacctct cctcctccca gttctcctct ggatcgcagt catccagaga 1380
tgtgacctcc tccagccgcc aaatccgcac caaggtcatg gatgtgcacg atggcaaggt 1440
ggtgtccacc cacgagcagg tccttcgcac caagaactga ggctgcccag ccccgctcag 1500
gcctaggagg ccccccgtgt ggacacagat cccactggaa gatcccctct cctgcccaag 1560
cacttcacag ctggaccctg cttcaccctc accccctcct ggcaatcaat acagcttcat 1620
tatctgagtt gcat 1634

<210> 362

<211> 1688

<212> DNA

<213> Homo sapiens

<400> 362
acagcacgct ctcagccttc ctgagcacct ttccttcttt cagccaactg ctcactcgct 60
cacctccctc cttggcacca tgaccacctg cagccgccag ttcacctcct ccagctccat 120
gaagggctcc tgcggcatcg gaggcggcat cgggggcggc tccagccgca tctcctccgt 180
cctggccgga gggtcctgcc gtgcccccag cacctacggg ggcggcctgt ctgtctcctc 240
tcgcttctcc tctgggggag cctgcgggct ggggggcggc tatggcggtg gcttcagcag 300
cagcagcagc tttggtagtg gcttcggggg aggatatggt ggtggccttg gtgctggctt 360
cggtggtggc ttgggtgctg gctttggtgg tggttttgct ggtggtgatg ggcttctggt 420
gggcagtgag aaggtgacca tgcagaacct caatgaccgc ctggcctcct acctggacaa 480
ggtgcgtgct ctggaggagg ccaacgccga cctggaagtg aagatccgtg actggtacca 540
gaggcagcgg cccagtgaga tcaaagacta cagtccctac ttcaagacca tcgaggacct 600
gaggaacaag atcattgcgg ccaccattga gaatgcgcag cccattttgc agattgacaa 660
tgccaggctg gcagccgatg acttcaggac caagtatgag cacgaactgg ccctgcggca 720
gactgtggag gccgacgtca atggcctgcg ccgggtgttg gatgagctga ccctggccag 780
gactgacctg gagatgcaga tcgaaggcct gaaggaggag ctggcctacc tgaggaagaa 840
ccacgaggag gagatgcttg ctctgagagg tcagaccggc ggagatgtga acgtggagat 900
ggatgctgca cctggcgtgg acctgagccg catcctgaat gagatgcgtg accagtacga 960
gcagatggca gagaaaaacc gcagagacgc tgagacctgg ttcctgagca agaccgagga 1020
gctgaacaaa gaagtggcct ccaacagcga actggtacag agcagccgca gtgaggtgac 1080
ggagctccgg agggtgctcc agggcctgga gattgagctg cagtcccagc tcagcatgaa 1140
agcatccctg gagaacagcc tggaggagac caaaggccgc tactgcatgc agctgtccca 1200
gatccaggga ctgattggca gtgtggagga gcagctggcc cagctacgct gtgagatgga 1260
gcagcagagc caggagtacc agatcttgct ggatgtgaag acgcggctgg agcaggagat 1320
tgccacctac cgccgcctgc tggagggcga ggatgcccac ctttcctccc agcaagcatc 1380
tggccaatcc tattcttccc gcgaggtctt cacctcctcc tcgtcctctt cgagccgtca 1440
gacccggccc atcctcaagg agcagagctc atccagcttc agccagggcc agagctccta 1500
gaactgagct gcctctacca cagcctcctg cccaccagct ggcctcacct cctgaaggcc 1560

accagcggct catggacatc aagtcgcggc tggagcagga gattgccacc taccgcagcc 1200
tgctcgaggg acaggaagat cactacaaca atttgtctgc ctccaaggtc ctctgaggca 1260
gcaggctctg ggcttctgc tgtcctttgg agggtgtctt ctgggtagag ggatgggaag 1320
gaagggaccc ttacccccgg ctcttctcct gacctgccaa taaaatttta tggtccaagg 1380
gaaaaaaaaa aaaaaaaaaa aaaaaaa 1407

<210> 360

<211> 2352

<212> DNA

<213> Homo sapiens

<400> 360
ccctgcactt gggagccggt agcactccta tcactgcttc tcaacccgtg agctaccagc 60
tgtgtcatga gctgcagaca gttctcctcg tcctacttga ccagcggcgg gggtggcggg 120
ggcggcctgg gcagcggggg cagcataagg tcttcctaca gccgcttcag ctcctcaggg 180
ggccgtggag gagggggccg attcagctct tctagtggct atggtggggg aagctctcgt 240
gtctgtggga ggggaggcgg tggcagtttt ggctacagct acggcggagg atctgggggt 300
ggttttagtg ccagtagttt aggcggtggc tttgggggtg gttccagagg ttttggtggt 360
gcttctggag gaggctatag tagttctggg ggttttggag gtggctttgg tggtggttct 420
ggaggtggct ttggtggtgg ctatgggagt gggtttgggg ggttaggggg ctttggaggt 480
ggtgctggag gaggtgatgg tggtattctg actgctaatg agaagagcac catgcaggaa 540
ctcaattctc ggctggcctc ttacttggat aaggtgcagg ctctagagga ggccaacaac 600
gacctggaga ataagatcca ggattggtac gacaagaagg gacctgctgc tatccagaag 660
aactactccc cttattataa cactattgat gatctcaagg accagattgt ggacctgaca 720
gtgggcaaca acaaaactct cctggacatt gacaacactc gcatgacact ggatgacttc 780
aggataaagt ttgagatgga gcaaaacctg cggcaaggag tggatgctga catcaatggc 840
ctgcggcagg tgctggacaa tctgaccatg gagaagtctg acctggagat gcagtatgag 900
actctgcagg aggagctgat ggccctcaag aagaatcata aggaggagat gagtcagctg 960
actgggcaga acagtggaga tgtcaatgtg gagataaacg ttgctcctgg caaagatctc 1020
accaagaccc tcaatgacat gcgtcaggag tatgagcagc tcattgctaa gaacagaaag 1080
gacatcgaga atcaatatga gactcagata acccagatcg agcatgaggt atccagtagt 1140
ggtcaggagg tgcagtccag tgccaaggag gtgacccagc tccggcacgg tgtccaggag 1200
ttggagattg agctgcagtc tcagctcagc aagaaagcag ctctggagaa gagcttggaa 1260
gacacgaaga accgctactg tggccagctg cagatgatcc aggagcagat cagtaacttg 1320
gaggcccaga tcactgacgt ccggcaagag atcgagtgcc agaatcagga atacagcctt 1380
ctgctcagca ttaagatgcg gctggagaag gaaatcgaga cctaccacaa cctccttgag 1440
ggaggccagg aagactttga atcctccgga gctggaaaaa ttggccttgg aggtcgagga 1500
ggaagtggag gcagttatgg aagaggatcc aggggaggaa gtggaggcag ctatggtgga 1560
ggaggaagtg gaggtggcta tggtggagga agtgggtcca ggggaggaag tggaggcagc 1620
tacggtggag gaagtggttc tggaggaggt agtggaggtg gctatggtgg aggaagtgga 1680
ggtggccata gcggaggaag tggaggtggt catagtggag gaagtggggg caactatgga 1740
ggaggaagtg gctctggagg aggaagtggg ggtggctatg gtggaggaag tgggtccagg 1800
ggaggaagtg gaggcagcca tggtggagga agtggtttg gaggtgaaag tggaggcagc 1860
tacggaggcg gtgaagaagc gagtggaagt ggtggcggct acggaggagg aagcggaaaa 1920
tcatcccatt cctagtcttc ttcctcaaaa tctggtgacc aagatgagac aaaaggcttc 1980
ctttcgcgat actagagcct ctgtaaactt ctcctgccca gcccccagct gagcatcccc 2040
agatgggcag actcccgatg aagcctgctt attggatcct gacatcagga atagctggca 2100
acaagcccct gtgcccgttt gggactgaac tgacttggca tatgtgctct ggctttccat 2160
tggcttccaa ccctgcttgg ctttggatgc cttcagggtt cccgagacag acctctttcc 2220
tctctctggc ctgggagctc tcacaccctg cacgatcttg actataataa agcttcccta 2280
ctgcaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2340
aaaaaaaaaa aa 2352

<210> 361

<211> 1634

<212> DNA

<213> Homo sapiens

<212> DNA

<213> Homo sapiens

<400> 358
ggtacctcct gccagcatct cttgggtttg ctgagaactc acgggctcca gctacctggc 60
catgaccacc acatttctgc aaacttcttc ctccaccttt gggggtggct caacccgagg 120
gggttccctc ctggctgggg gaggtggctt tggtgggggg agtctctctg ggggaggtgg 180
aagccgaagt atctcagctt cttctgctag gtttgtctct tcagggtcag gaggaggata 240
tgggggtggc atgagggtct gtggctttgg tggagggggct ggtagtgttt tcggtggagg 300
ctttggaggg ggcgttggtg ggggttttgg tggtggcttt ggtggtggcg atggtggtct 360
cctctctggc aatgagaaaa ttaccatgca gaacctcaat gaccgcctgg cctcctacct 420
ggacaaggta cgtgccctgg aggaggccaa tgctgacctg gaggtgaaga tccatgactg 480
gtaccagaag cagaccccaa ccagcccaga atgcgactac agccaatact tcaagaccat 540
tgaagagctc cgggacaaga tcatggccac caccatcgac aactcccggg tcatcctgga 600
gatcgacaat gccaggctgg ctgcggacga cttcaggctc aagtatgaga atgagctggc 660
cctgcgccag ggcgttgagg ctgacatcaa cggcttgcgc cgagtcctgg atgagctgac 720
cctggccagg actgacctgg agatgcagat cgagggcctg aatgaggagc tagcctacct 780
gaagaagaac cacgaagagg agatgaagga gttcagcagc cagctggccg gccaggtcaa 840
tgtggagatg gacgcagcac cgggtgtgga cctgacccgt gtgctggcag agatgaggga 900
gcagtacgag gccatggcgg agaagaaccg ccgggatgtc gaggcctggt tcttcagcaa 960
gactgaggag ctgaacaaag aggtggcctc caacacagaa atgatccaga ccagcaagac 1020
ggagatcaca gacctgagac gcacgatgca ggagctggag atcgagctgc agtcccagct 1080
cagcatgaaa gctgggctgg agaactcact ggccgagaca gagtgccgct atgccacgca 1140
gctgcagcag atccaggggc tcattggtgg cctggaggcc cagctgagtg agctccgatg 1200
cgagatggag gctcagaacc aggagtacaa gatgctgctt gacataaaga cacggctgga 1260
gcaggagatc gctacttacc gcagcctgct cgagggccag gatgccaaga tggctggcat 1320
tggcatcagg gaagcctctt caggaggtgg tggtagcagc agcaatttcc acatcaatgt 1380
agaagagtca gtggatggac aggtggtttc ttcccacaag agagaaatct aagtgtctat 1440
tgcaggagaa acgtcccttg ccactcccca ctctcatcag gccaagtgga ggactggcca 1500
gagggcctgc acatgcaaac tccagtccct gccttcagag agctgaaaag ggtccctcgg 1560
tcttttattt cagggctttg catgcgctct attcccctc tgcctctccc caccttcttt 1620
ggagcaagga gatgcagctg tattgtgtaa caagctcatt tgtacagtgt ctgttcatgt 1680
aataaagaat tacttttcct tttgcaaat 1709

<210> 359

<211> 1407

<212> DNA

<213> Homo sapiens

<400> 359
cgcgaatcgc agcttctgag accagggttg ctccgtccgt gctccgcctc gccatgactt 60
cctacagcta tcgccagtcg tcggccacgt cgtccttcgg aggcctgggc ggcggctccg 120
tgcgttttgg gccgggggtc gcctttcgcg cgcccagcat tcacgggggc tccggcggcc 180
gcggcgtatc cgtgtcctcc gcccgctttg tgtcctcgtc ctcctcgggg gcctacggcg 240
gcggctacgg cggcgtcctg accgcgtccg acgggctgct ggcgggcaac gagaagctaa 300
ccatgcagaa cctcaacgac cgcctggcct cctacctgga caaggtgcgc gccctggagg 360
cggccaacgg cgagctagag gtgaagatcc gcgactggta ccagaagcag gggcctgggc 420
cctcccgcga ctacagccac tactacacga ccatccagga cctgcgggac aagattcttg 480
gtgccaccat tgagaactcc aggattgtcc tgcagatcga caatgcccgt ctggctgcag 540
atgacttccg aaccaagttt gagacggaac aggctctgcg catgagcgtg gaggccgaca 600
tcaacggcct gcgcagggtg ctggatgagc tgaccctggc caggaccgac ctggagatgc 660
agatcgaagg cctgaaggaa gagctggcct acctgaagaa gaaccatgag gaggaaatca 720
gtacgctgag ggccaagtg ggaggccagg tcagtgtgga ggtggattcc gctccgggca 780
ccgatctcgc caagatcctg agtgacatgc gaagccaata tgaggtcatg gccgagcaga 840
accggaagga tgctgaagcc tggttcacca gccggactga agaattgaac cgggaggtcg 900
ctggccacac ggagcagctc cagatgagca ggtccgaggt tactgacctg cggcgcaccc 960
ttcagggtct tgagattgag ctgcagtcac agctgagcat gaaagctgcc ttggaagaca 1020
cactggcaga aacggaggcg cgctttggag cccagctggc gcatatccag gcgctgatca 1080
gcggtattga agcccagctg ggcgatgtgc gagctgatag tgagcggcag aatcaggagt 1140

agtctcgccg gtgctgcagg gtacatctct tctgctaggt cgggcctctc tggccttggg 180
agctgcttgc ctggctccta cctgtcttct gagtgccaca cctctggctt tgtggggagc 240
gggggctggt tctgcgaggg ctccttcaac ggcagcgaga aggagactat gcagttcctg 300
aacgaccgcc tggccaacta cctggagaag gtgcgtcagc tggagcggga gaacgcggag 360
ctggagagcc gcatccagga gtggtacgag tttcagatcc catacatctg cccagactac 420
cagtcctact tcaagaccat cgaagatttc cagcagaaga tcctgctgac taagtctgag 480
aatgccaggc tggtcctgca gattgataat gccaagctgg ctgctgacga cttccggacc 540
aagtatgaga cagagctgtc tctgcggcag ctagtggagg ccgacatcaa cggcctgcgt 600
aggatcctgg atgagctgac cctgtgcaag gctgacctgg aggctcaggt ggagtccctg 660
aaggaggagc tgatgtgcct caagaagaat cacgaggagg aagtcagtgt actccgttgc 720
caacttgggg accgactgaa tgtggaggtg gacgctgctc cccagtgga tctcaacaag 780
atcctggagg atatgagatg ccagtacgag gccctggtgg agaataaccg cagagatgtg 840
gaggcctggt tcaacaccca gactgaggag ctgaaccagc aggtggtgtc cagctcggag 900
cagctgcagt gctgccagac ggagatcatc gagctgagac gtacggtcaa cgcgctagag 960
attgagctgc aggctcagca cagcatgcgg aattccttgg aatccaccct ggccgaaacc 1020
gaggcccgct acagctccca gctggcccag atgcagtgcc tgatcagcaa cgtggaggcc 1080
cagctgtctg agatccgctg cgacctggag cggcagaacc aggagtacca ggtgttactg 1140
gacgtcaagg cccggctgga gggcgagatc gctacctacc gccacctgct ggagggagag 1200
gactgcaagc ttcctcccca accttgtgcc acggcatgca agcctgttat tagagttcct 1260
tctgtccccc cggtgccctg tgtcccctct gtgccctgca ccccggctcc ccaggttggc 1320
actcagatcc gcaccatcac cgaggagatc agagatggga aagtcatctc ctccagggag 1380
cacgtgcagt cccgcccgct gtga 1404

<210> 357

<211> 1693

<212> DNA

<213> Homo sapiens

<400> 357
cacagtcctc ggcccaggcc aagcaagctt ctatctgcac ctgctctcaa tcctgctctc 60
accatgagcc tccgcctgca gagctcctct gccagctatg gaggtggttt cggggggtggc 120
tcttgccagc tgggaggagg ccgtggtgtc tctacctgtt caactcggtt tgtgtctggg 180
ggatcagctg gggctatgg aggcggcgtg agctgtggtt ttggtggagg ggctggtagt 240
ggctttggag gtggctatgg aggtggcctt ggaggtggct atggaggtgg ccttggaggt 300
ggctttggtg ggggttttgc tggtggcttt gttgactttg gtgcttgtga tggcggcctc 360
ctcactggca atgagaagat caccatgcag aacctcaacg accgcctggc ttcctacctg 420
gagaaggtgc gcgccctgga ggaggccaac gctgacctgg aggtgaagat ccgtgactgg 480
cacctgaagc agagcccagc tagccctgag cgggactaca gcccctacta caagaccatt 540
gaagagctcc gggacaagat cctgaccgcc accattgaaa acaaccgggt catcctggag 600
attgacaatg ccaggctggc tgtggacgac ttcaggctca agtatgagaa tgagctggcc 660
ctgcgccaga gcgtggaggc cgacatcaac ggcctgcgcc gggtgctgga tgagctcact 720
ctgtctaaga ctgacctgga gatgcagatc gagagcctga atgaagagct agcctacatg 780
aagaagaacc atgaagagga gatgaaggaa tttagcaacc aggtggtcgg ccaggtcaac 840
gtggagatgg atgccacccc aggcattgac ctgacccgcg tgctggcaga gatgagggag 900
cagtacgagg ccatggcaga gaggaaccgc cgggatgctg aggaatggtt ccacgccaag 960
agtgcagagc tgaacaagga ggtgtctacc aacactgcca tgattcagac cagcaagaca 1020
gagatcacgg agctcaggcg cacgctccaa ggcctggaga ttgagctgca gtcccagctg 1080
agcatgaaag cggggctgga gaacacggtg gcagagacgg agtgccgcta tgccctgcag 1140
ctgcagcaga tccagggact catcagcagc atcgaggccc agctgagcga gctccgcagt 1200
gagatggagt gccagaacca agagtacaag atgctgctgg acatcaagac acgtctggag 1260
caggagatcg ccacctaccg cagcctgctc gagggccagg acgccaagaa gcgtcagccc 1320
ccgtagcacc tctgttacca cgacttctag tgcctctgtt accaccacct ctaatgcctc 1380
tggtcgccgc acttctgatg tccgtaggcc ttaaatctgc ctggcgtccc ctcccctctgt 1440
cttcagcacc cagaggagga gagagccggc agttccctgc aggagagagg aggggctgct 1500
ggacccaagg ctcagtccct ctgctctcag gaccccctgt cctgactctc tcctgatggt 1560
gggccctctg tgctcttctc ttccggtcgg atctctctcc tctctgacct ggatacgctt 1620
tggtttctca acttctctac cccaaagaaa agattattca ataaagtttc ctgcctttct 1680
gcaaacataa aaa 1693

<210> 358

<211> 1709

tgtgccctcc ccatgcgtga cccgcaccgt ctgtgtgcca cgcactgttg gcatgccttg 1380
ctcaccctgc ccccagggcc gctactgaag tccctttgtg ccagtggatc ctggagggcc 1440
tggggctggg cagcctggta ttcagtggcc accagaagag cagggccagc cccggtcagc 1500
aaggaagacc ctgagcagga ccgtggatca cctgcaacaa gctctgatac tccggtcagc 1560
acttaagccc tcatcacttc aaaactgcct ctttttccca tgggtgaact gttctctttg 1620
gtgatgtttc tggttgtctg tgctgcctca aagagcgtgt gttcttagtt aactggcaaa 1680
tagagctgta ctcagtggcc ttgcaaac 1708

<210> 355

<211> 2051

<212> DNA

<213> Homo sapiens

<400> 355

tgggctcaaa tgacaggtcc ctttgaagaa ctgacaggga cttggtcttc ccagtgatgt 60
ggggtggtgg agatgagtct tgacatagta atggggctcc tattttatga gactggtaaa 120
ttttcacagg aaactaaata acaggttaat gaggggtttt aatagaaagt ggtgttaact 180
ttaatgagga gaaaacactt tgtaaaggcc aaccaatgcc cactggggca gagatataaa 240
tctggggtaa ggagctgctt ctctgctttg ggtgtctgg cctcagagac ctatcaattg 300
catctgagtt gcaggggcca tggcttccaa atgcctcaag gccggcttct cttctgggtc 360
tctcaagagc ccaggagggg ccagtggggg ctccactcgt gtgtccgcaa tgtactccag 420
cagctcttgc aagcttccaa gtctctcccc tgtggccaga agtttctctg cctgctcagt 480
gggtctgggc agaagcagct acagggccac cagctgcctc cctgctctct gcctccctgc 540
tggaggcttc gctaccagct acagtggggg tgggggctgg tttggggagg gcatcctcac 600
tggcaatgag aaggagacca tgcaatccct gaacgaccgc ctggccggct acctggagaa 660
ggtgcgtcag ctggagcagg agaacgccag cctggagagc cgcatccgtg agtggtgtga 720
gcagcaggtc ccctacatgt gccctgacta ccagtcctac ttccggacca tcgaggagct 780
ccagaagaag actctatgca gcaaggctga gaatgccagg ctggtggtgg agattgacaa 840
tgccaaattg gctgcagatg acttcaggac caagtatgag acggaggtgt ccctgcggca 900
gctggtggag tcagacatca acggcctgcg caggatcctg gatgacctga ccctgtgcaa 960
gtctgacctg gaggcccagg tggagtccct gaaggaggag ctgctctgcc tgaagaagaa 1020
ccatgaggag gaagtgaact cactgcgctg ccaacttggt gaccgcctca atgttgaggt 1080
ggacgctgcc cccacctgtg acctgaaccg agttctggag gagatgaggt gccagtatga 1140
aaccctggtg gagaataacc gccgggatgc tgaagactgg ttggacaccc agagtgagga 1200
gctgaaccag caggtggtgt ccagctcaga gcagttgcag tcctgccagg cagagatcat 1260
cgagctgaga cgcacggtca acgccctgga gattgagctg caggctcagc acagcatgag 1320
agatgctttg gaatccaccc tggcagagac ggaggcccgc tatagctccc agctggccca 1380
gatgcagtgc atgatcacca acgtggaggc ccagctggcc gagatccggg ctgacctgga 1440
gcggcagaac caggagtacc aggtgctgct ggacgtccgg gcccggctgg agtgtgagat 1500
caacacgtac cggggcctgc tggagagtga ggacagcaag ctcccctgta acccatgtgc 1560
acctgactac tcaccctcca agtcatgcct tccctgtctt cctgcggcct cctgcggtcc 1620
tagtgcagcc cgcacaaact gcagcccccg ccccatttgt gtgccctgcc caggggtcg 1680
gttctgagag cgggtgaccc agatggccat ggctattgtc tccagggctt gaacttggcc 1740
tctacccaaa cttaaccctt gtagcccaat cccctctctt cgcgcagagc ccaggcccag 1800
ggtctggctg aaaaggcttt ctgcaaatac atgccctaaa gtttctcaga gcctgtcaca 1860
aaggccggct gcccccaaag gtctcaactc ctcatcattt caatgggtgc cagggtctct 1920
gttctcaggc tgcctcctgg gtcaggtttt ccttctaggt gctgttccgg tggattctga 1980
aatgcagtag agggcttttg ttggcagaac aataaagtgc atttgctcag gcccctgatg 2040
cctaacttgc a 2051

<210> 356

<211> 1404

<212> DNA

<213> Homo sapiens

<400> 356

atggccaccc agacctgcac ccctaccttc tccactgggt ctatcaaggg cctctgtggc 60
acagcaggcg gcatctctcg ggtgtcctcc atccgttctg tgggctcctg caggtcccc 120

gagctccaac agaagatcct gtgcagcaag gccgagaatg ccaggctgat tgtacaaatt 960
gacaatgcca agctggctgc cgatgacttt aggatcaagc tggagagtga gcgctccctg 1020
cgccagctgg tggaggcaga caagtgtggg acacagaagc tcctggatga tgcgaccctg 1080
gccaaggccg acctggaggc ccagcaggag tccctgaagg aggagcagct ctccctcaag 1140
agcaaccacg agcaggaagt aaagattctg aggagtcagc tgggggagaa gctccggatt 1200
gagctggaca ttgagcccac cattgacctg aacagggtgc tgggggagat gcgggctcag 1260
tatgaggcca tgttggagac caaccgccag gatgtggaac agtggttcca agcccagtct 1320
gaaggcatca gcctgcagga catgtcctgc tccgaggagc tgcagtgctg ccagtcggag 1380
atcctggagc tgagatgcac ggtgaatgcc ctggaggtgg agcgccaagc ccagcacacc 1440
ttgaaggact gtctgcagaa ctccctgtgt gaagccgagg accgcttcgg cacggagctg 1500
gcccagatgc agagcctcat cagcaacgtg gaggagcagc tgtctgagat ccgggccgac 1560
ctggagcggc aaaaccagga gtaccaggtg ctgctggacg tgaagacccg gctggagaat 1620
gagattgcca cgtaccggaa ccttctggaa agcgaggact gcaaactccc ctgcaatccg 1680
tgctccacgt ctccctcctg cgtgactgcc ccctgtgctc ctcgcccaag ctgtggcccc 1740
tgcaccacct gtgggcccac ctgtggagcc agcaccaccg gaagccgatt ctgaattcct 1800
gtggacccac aggggctggc taaggcgagg gatacccaaa gagagatgct tgttatacct 1860
ttagaaaatc tggcttctaa ctttctgtat gtataggcct gtccaaaggc tatgagatac 1920
cagggacagt ggaatcttga tggaaatcct tcctttcctg ctcttggttt tcccaggtga 1980
gctctatgcc ctcagtggct ggattgcagc cacccatggt ggctcagaat atctggattg 2040
gagaggcata aggttgaaaa gcattttgga gtaggaacac ggcatgattt aaaattgcac 2100
ttgcacttca agaacaccaa ctttaacttg aaatttttat agtttatgac aaaggaagtg 2160
aagaacagac ttcactgtct gagatttccc acatttcttg ggttgtttct aacattggca 2220
gcctggatta tggccctagg gagagggaaa tagaatcaga attgtacctg accctgtctt 2280
ttctaattcc cttctcctat cctcctgcct ccctatcttg gaaactgatt aagtggaaac 2340
tttccttgta tcaaatcctg agttttctct acttcagggt ttctagcttt gacaaaggac 2400
caggcttgcc ttttctgtgg caacacagcc tcttacatat ggcttccata gcttcttgct 2460
atgtgaaagg aactttcaga ttaagaaatt tctctctttt ttctctataa attctgactc 2520
tccagacttg tctggctata acccctgtcc ctgtctaatt ctcccagtct aaaatctaaa 2580
caatgacttc atttttttct agtactttct cctgaagtat tgaaacctat tgattcaatg 2640
ggtagagatt tgctaagtat gatgtgcttc cacctcttct cttcatactc tctaccttcc 2700
ttaccatgat tcagccacac tggcctttct gttccttgga catgcctgga gacctgtccc 2760
tgctcttcct tctctctggg gtgctcttcc ccctccttat aaatgttcag gtcagcggaa 2820
atgttccttc ctctgag 2837

<210> 354

<211> 1708

<212> DNA

<213> Homo sapiens

<400> 354
ggcggcatgg ggtctgtaga catccaggta gctgtggctg aggagaaagg gcctctccaa 60
catgacatcc tcctgctgtg tcaccaacaa cttgcaagcc tctctcaaga gctgcccccg 120
gcctgcctcg gtctgttcca gcggcgtgaa ctgccggcct gagctgtgcc tgggctatgt 180
ctgccagccc atggcatgcc tgccttcggt ctgcctgccc accaccttcc ggccagccag 240
ctgcctctcc aaaacctatc tatccagttc ctgccaggca gccagtggca tctccggctc 300
catgggcccc ggcagctggt acagcgaagg ggccttcaat ggcaatgaga aggaaaccat 360
gcagttcctt aacgaccgcc tggccagcta cctgacgagg gtgcggcagc tggagcagga 420
gaatgcggag ctggagagca ggatccaaga ggcctctcac tcccaggtgc tcaccatgac 480
tcctgactac cagtctcatt tcaggaccat tgaggagctc cagcagaaga ttctgtgtac 540
caaggcagag aatgccagga tggttgtgaa cattgataat gccaaactgg ctgccgatga 600
cttcagggcc aagtacgagg cagagctggc catgcggcag ctggtggagg ccgacatcaa 660
tggcctgcgc aggatcctgg atgatctcac tctgtgcaag gctgacctgg aggcccaggt 720
tgagtccctg aaggaggagc tgatgtgcct caaaaagaac catgaggagg aagtcggttc 780
ccttcgatgc cagcttgggg accgccttaa catcgaggtg gacgctgcac ccccggtgga 840
cctgaccagg gtgctggagg agatgcggtg tcagtacgag gccatggtgg aggccaaccg 900
cagggacgtg gaggaatggt tcaatatgca gatggaggag cttaaccaac aggtggccac 960
aagctctgag cagcttcaga actaccagtc agacatcatt gacctgagac gcacggtcaa 1020
cacgctggag atcgagctgc aggcccagca cagcctgagg gactccctgg aaaacacgct 1080
gacggagagt gaggcccgct acagctccca gctggcccag atgcagtgca tgatcaccaa 1140
cgttgaggcc cagctggctg agatccgggc tgacctggag cggcagaacc aggagtacca 1200
ggtgctgctg gacgtccggg cccggctgga gggcgagatc aacacgtacc ggagcctgct 1260
ggagagcgag gactgcaagc tgccctgtaa cccatgctcc actccttcct gcaccacctg 1320

<213> Homo sapiens

<400> 352
tggggcctga atcaactgct ggtgtttgtt gagcagaaac agtgagtctc tgacttccca
ccaacagcta acctgatatg gaagagccaa gaggcatctc agatttgtca acattgcttt 60
aatgagctga ggaacttctc atggtaaact cagcagctga gtaacaggat gatggcacca 120
caacaggtag ataatcaagg caggaggtca aagcactgga gccaacaccc gcccaggatg 180
gggtataaaa gggttgggag gagaggaggc ttcagtctca gtggctcagc cttcccagct 240
gatctgaagc tcctgtgcag cctcagccct acaccatgac ctccttctac agcacctcct 300
catgccctct gggttgcacc atggctcctg gagcaagaaa tgtctttgtc tctcctatcg 360
atgttgggtg ccagcctgtg gcagaggcca atgctgcctc catgtgcctc ttggccaacg 420
tggcacacgc caacagagtc cgtgtggggt cgactcccct gggccgcccc agcctctgtc 480
tgcccccaac cagtcacact gcttgtccct tgccagggac ctgtcacatt cccggcaaca 540
tcggaatctg tggggcctac ggcaaaaaca ccctgaatgg ccatgagaag gagaccatga 600
agttcctgaa tgaccgcctg gccaactacc tggagaaggt gcgccagctg gagcaggaga 660
atgcagagct ggagaccaca ctcctcgaga ggagcaagtg ccacgagtcc accgtgtgcc 720
ccgactacca gtcctacttc cgtacaatcg aggagctcca gcagaagatc ctgtgcagca 780
aggctgagaa tgccaggctg attgtacaaa ttgacaacgc gaagctggct gctgatgact 840
ttaggatcaa gctggagagt gagcgctccc ttcaccagct ggtggaggcg gacaagtgcg 900
ggacgcagaa gctcctggat gacgcgaccc tggccaaggc cgacctggag gcccagcagg 960
agtccctgaa ggaggagcag ctctccctca agagcaacca cgagcaggaa gtaaagattc 1020
tgaggagtca gctgggggag aagttccgga tcgagctgga cattgagccc accattgacc 1080
tgaacagggt gttgggggag atgcgggctc agtacgaggc catggtggag accaaccacc 1140
aggatgtgga acagtggttc caagcccagt ctgaaggcat cagcctgcag gccatgtcct 1200
gctccgagga gctgcagtgc tgccagtcgg agatcctgga gctgagatgc acggtgaatg 1260
ccctggaggt ggagcgccaa gcccagcaca ccttgaagga ctgtctgcag aactccctgt 1320
gtgaagcgga ggaccgctac ggcacagagc tggcccagat gcagagcctc attagcaact 1380
tggaagagca gttgtctgag atccgggccg acctggagcg gcagaaccag gagtaccagg 1440
tgctgctgga cgtgaaggcc cggttggaga acgagattgc cacataccgg aaccttactc 1500
ccctgcaatc cctgttccac gcctgcctcc tgtacttctt gtccaagctg tggccctgtc 1560
accggtgggt ctccctctgg ccatggagcc agcatgggga gatgattctg aaggcccgag 1620
ttaggagatt gaggctggtt gcactggggt caggagtgcc ctcaccttgc ccagtcttcc 1680
ttcaagacta gactcactga ggcattttcc ctaaatcaac cggtagcaga tacttccaag 1740
gagtggctcc ctgcctatgc ctctaagctg tatttttgt tgttgctaaa atgttgtaat 1800
tatcttccat gaaagcaatt atttttcctg gtgtctcttc ttggtacttt tagttctatt 1860
ccagtgtttc agaatctcca aaaatgtaac tgggttccct gcattaaatg gtcaataaac 1920
ctcctt 1980
<210> 353 1986

<211> 2837

<212> DNA

<213> Homo sapiens

<400> 353
tagtcctcct gctacaccac tgaccaacag gaaagtttgt gtctccagag tggacacatc
cataaagagg ccaaacccag tcaaggtcta agcatctgat ggctataact ttgcttcttt 60
gaaagataat aaaaagcttc tgacctccca tcaacagcca acctgatttc aaagagccaa 120
gaggcctcag attcgtcaac attcctttaa cgagctgaga aacttttcat ggtaaactca 180
gcagctgagt aacaggatga tggcaccaca acaggtagat aatcaaggca ggaggtcaaa 240
gcattggagc caacacccgc ccaggatggg gtataaaagg gctgggagga gaggaggctt 300
cagtctcagt ggttcagctt tcccagctga tctgaagctc ctgtgcagcc tcagcccaac 360
accatgacct cttcctacag cagctcctca tgccctctgg gttgcaccat ggctcctgga 420
gcaagaaatg tctctgtctc tcccatcgac attgggtgcc agcctggggc agaggccaac 480
attgccccca tgtgcctttt ggccaacgtg gcacatgcca accgagtccg tgtggggtcc 540
actcccctgg gccgccccag cctctgtctg ccgcctacct gccacactgc ttgtcccttg 600
ccagggacct gccacattcc tggcaacatt ggaatctgtg ggcctatgg tgaaaacacc 660
ctgaatggcc atgagaagga gaccatgcag ttcctgaatg accgcctggc caactacctg 720
gagaaggtgc gccagctgga gcaggagaat gcggagctgg aggccacact cctcgagagg 780
agcaagtgcc acgagtccac cgtgtgcccc gactaccagt cttacttcca caccatcgag 840
900

agtttacaac atgtgctaga atgcatcatc cacctggtaa gcaacttgtg aagcaggcgg 3120
gcagggatag aatctcctcc acaggtgggg accctgaggc ttggggtgct ggtgtgactt 3180
gccagaagcc atgtgggtac cgagtggcag agctgggact tggccttaaa ttttgtacca 3240
aattttcctt tcttgaagca atttcctgtg aaactgttgc cttaataagg atatgtttct 3300
gtgctgcagc ttgattaatt tgtgagtgct gataattcat agcagaaatc atgctccata 3360
gcctaagatg aaaacatgtg aaagtcaaag ggttgacacc acctgaggag ggtctaccct 3420
gggtgctgga agggcattgt gttttatcat ctttcttcgg ggcactgtgc aagtctcact 3480
cttctagagc taaagaaagc catttccttc ctctcatgct ggcaatgtgg cttttactct 3540
aataggaat ccagggtccc tttgggtatg gggcttgtag agaggggcct gcagtgattg 3600
cactgtgttg caaatgcctg gttactttgc cacctccctt ttgggccatt gtgtgtaaag 3660
gcaggtgctg gctgtgtgtc catcacagta tcctcagtac cagctgctcc ccttggttca 3720
tctcaggtgc ttagtattga ctgagtgggg aacatcctgt tgctttctgg gctttgagca 3780
cactacacag gataagccaa ctcacaatta gatggcaaag ggcaggccag tggacttaca 3840
gctgcgagct gggagtcatg tcctggtgcc ccttcctgcc ctctaactgt cactcctgca 3900
ctttccctgt tacaggtacg agagagagcg ctcgctgcaa cagctggtgg aggctgacat 3960
ctgtggcctg cgcagggtgc tagacaacct caccctcgcc aagtgtgacc tggaggccca 4020
gctggagtcc ctgaaggagg agctgctttg cctcaagaag aaccatgagc aggtgtgtga 4080
cccttggtgt gatgcattga gactttctgg aggtgaattc aggctcatct gggggaaaa 4140
aacggaggct tccttctgct tgggattgaa gctttgcagg tgagctctgg cttttggtgg 4200
gagattgctt ctcatccaca gcttctggaa gctgctaggc tggtgcaggc agcggtttat 4260
cttcagagga gcacagggaa attactgggg aggatggaaa gtgggagttg agagtttggg 4320
tgaagggcag ttgggggccc tgcaggcggg tcatactcct ctgatgccta caggaagccc 4380
acactctaag ggtcagctg ggagacaagc tccggataga gctggacatt gagcccacca 4440
ttgacctgag tagagttctg gggagacgt gaggccagta cgaggccatg gtggagacca 4500
atcgccagga tgtggagcag tggttccaag cccaggtgag ggggactgct aagcagtggg 4560
gattggcctc catctgggca ggtaggtgac catgtttctc tatgctccag tctgaaggca 4620
tcagcctgca ggccatgtcc tgctctgagg agctgcagtg ctgccagtca gagatcctgg 4680
agttgagacg ctcggtgaac gccctggagg tggagcttca ggctcagcac acactggtga 4740
gtccctctgc acgcatggga gacccagctt ggctcccact acccatgagt agctctctgg 4800
ggtcagtaca ggcctcgaat accaaccacg gactcagctt aaaccttggt ttcaactctt 4860
ttggaagcac ttccctgtag atgtggatgc tgagtgaatc aacattaact taggtcagag 4920
ggctctgatc ttaaaggtct accatgaagt gtgttagtta ttctggctga agacagctaa 4980
agggagttag attgctctct ggacctggca atgtcttggg gtgtcacaat cttaagaggt 5040
tggagatcca tgatttttaa cttgcttcct cctttctttg atattataga aggactgtct 5100
acagaactcc ctgtgtgaag ccgaggaccg ctaccgcaca gagctggccc agatgcagag 5160
cctcatcaac aatgtggagg aacagctgtc tgagatccgg gctgacctgg agcggcagaa 5220
ccaggagtac caggtgctgc tggatgtgaa ggcccggctg cagaatgaga ttgccacata 5280
ccggaacctt ctggagagtg aggactgcaa gtacgtaccc tgttgtctat gtcaacaccg 5340
tgtgcaaaaa cagataactt ctcttgggaa tgacctaaaa ggagctctgg aggccaggtg 5400
ggggtggcat tgcagatccc atctttggag acaaacttgc taccactagg cctggtgagc 5460
taatgcctgc ctaggcgtag gtgagggtgt acaggagttt tcaggtcatg gcatttgttt 5520
tgctgaccac gttgtgctcc cagttgctgg ctactcccat agagggctaa aatactgaat 5580
gcccagcaca ttcaggtttt gctaaattct ctccaggaga agtattccaa taacgagtat 5640
tctttgtgct catagcattt aaagtattgc taagggcatg tgcactgact ctttcatttt 5700
agttccacaa tagccctgtg aaaagcatca gagggctgat atacccccttt aagcactaag 576
gaacactatg aggggctgtg tttaaatgga atcagtggct gacatctgga ccattcttgt 5820
tcctctggcc tttctgcctt tctctctgct gtgttccttc catatgggac gtgaccctgg 5880
cctggcacac aggtacactg agttctggcc ccagctccac tgtggactgt gggcacgtca 5940
ttgccccttt caggcttaga gggacctcta aggtctaagg tcctcttttg cactttaatt 6000
ctaaagtgtc tataatatac ttttgtcttt cacatacttc tcatgccaat gtttcaaaca 6060
cctttttct ctatgcagat ttccctgcaa cccgtgtgca accccagcct tcagcactcc 6120
tagtccagcc cctgcagcct gcgcccccctg ctcccgggcc acccatgggc cctgctcatc 6180
aactggatac tgacacactc ctcagcccga ggtgctggga aaggggaggt gtcttcagct 6240
cgttaggctt tgctctgcag gcctgctaac ctgggccttg cctccctggc cagccaggca 6300
ggagactgaa gagagacagt gccactctgt ggataggtgt ttgaatccag cagctcagcc 6360
cctatggctg gatttgtttc ttcccctcaa ggtgtcccaa gtgctccatt ttctctcttc 6420
tgttgcctcc tgatctctct gagatggggc aggcactttt gtttttattt tgctattaaa 6480
cttcgcttct gcagcagaat gagcccagtg cagtccagcc tcagcatctc cgtcccatgc 6540
tgagagcctt gctgggaact ccttgagcca catgcttctg tagacggggg ttgtggctgg 6600
ctcgccagag tggttctggg agatggcaaa gccagtgaag gctagagtta aaaatggac 6660
tgcggccaaa gtg 6673

<210> 352

<211> 1986

<212> DNA

aggggcctga aaacatcatt tgcatggctg gagttgcccg cgtaaggcag ccaagaaact 1440
cacccaaagc ctgtagcctc cccaactact ccagactgtc ctgctcaccc tttccttcct 1500
gggggtctgt tccttcctat gctcacccag agaactctct gatgtgccag tgggcctccc 1560
ttttaacctc ctaataaata tcatttcctt ggcaaaaaaa aaaaaaaaaa aaaaaa 1616

<210> 351

<211> 6673

<212> DNA

<213> Homo sapiens

<400> 351

aagaccatgg tcctggtatt taagactttt cacaacctgg actcaatgca actttccggg 60
gcttggggat aatgtctcca ttataaatac caacacaact agactatttt ccattttcca 120
aatgtttgat tgtgtttatt ttgttctcta tgagagaagc tcattcccca accctttata 180
aaaaatatca ttcaaaacat atttggaaaa tcccctactc catgaagtct ttttccggtc 240
tctcagtgag ctctcttcca cctcgaaatc tttgcagctg tcttttaaaa aaaaaaaata 300
cttcctttta ctgtcttcat aatatcaaga tcatttactt ttttatcttt cttattaaaa 360
tctaaaatct ccttcttaca agacagcata gatgtgcgtt tgttccgcac tgggcactct 420
acaaatatct atcaaaacaa acaaatcatg atttttccct ttcttcaatg tgcctgattg 480
actaatattt taggttttgc tgcattggga acattcagag gtaggaaaag aatgtgggag 540
tgagagagac ttgcaggaca ttttaggagc ctttagttct tgattatttc atagacataa 600
gctgacagaa ggggaaagat ttcacagcca cagaaagtgg gctgcttgaa gaagaccttg 660
ttttctctgt tttccaggac cctgcagtga tgggaagaaa gaggttaggg cctctttcat 720
gagctcatta actctttata ggaaacaaaa accactgact aatagagtga ttgtataaca 780
attgacataa atattaatga ggaaaatgcc agtttttcttg caccaacacc cactaaggct 840
tatataaggt ctacaggaga aggaagattt tggtgtcaag cagatccaac tgtgggcaaa 900
gcaggatctt gcctttttagc accatgactt ctgaccattg cagttccctc ctcagcgggc 960
aggtttcaga ggccaacgct gcctctctgt gcctcttggc taatgtggca catgccaatc 1020
gagtccgtgt ggggtcgact cccctgggcc gcctcagcct ctgtctgccc ccaacctgcc 1080
acaccacttg tcccttgcca gggacctgcc acattcctgg caacatcgga atctgtgggg 1140
cctaccgcga aaacaccctg aacggccacg agaaggagac catgcagttc ctaaacgacc 1200
gcctggccaa ctacctggag aaggtgcgcc agctggagtg ggacaatgca gaactggaga 1260
ccaaactcca tgagaggagc aagtgccacg agtccagcgt gtgccgaaac taccagtcct 1320
acttctgcac catccaggag ctccagcaga aggtgaggtt tgcggtgcac cagatcaggg 1380
gtcaggaatc tgcttactgc ctgtcggcca aatccggccc accacctgct tttgcaaata 1440
aagttttatt ggttcatggc catgctcatg catttgtatg ttgtctgcag ctgcttttgt 1500
actacagtgg cagagttcag tcattgtaat ggagactgta gggtctgcaa agcctaaagt 1560
atttactatc tgacccttta caaaaaaaga ttactgaccc tgggtcagca gtggattttt 1620
tttttctttt aatgtgggaa gagaatctct tgagtgaagg tgaatgggaa tttgtaaagg 1680
atgagatgta gaatttatca ctctaggatt ctgagttcat gcttacttaa cttcagcaca 1740
taaaattcct aaatcgtagt cacaaaaatc atgctatttt catgaaaacc atgcaaaaac 1800
cgtgctgagt ttatgaaagt tgcagtatac tgaaactttc tggtgcagta aacagggcaa 1860
aaggaatatt tcctaacatc cttgccaagg cagtgtctat ccacacaaca gcatcaacct 1920
tcttatgcat attctattgt cacctttgat tatccagtta aagcttttga agagcctaca 1980
ttcttagttt tggagacaag aagacaagga atgtcaaaat aggggttaca gaaacctggc 2040
tagattcaga gaagacctgt ggggaggttg gacaaccaga ggaaagattc tttctttaga 2100
acagatccta ctttatctaa aacattcata gccagtcctc tgttctgtgg cctagttttt 2160
tttttttttt ttttgagaca aagtctcgct ctgtcaccca ggctggagtg cagtggtgca 2220
atctcagctc actgcaaact ctgcctccca ggttcaagtg attctcatgt ctcagcttcc 2280
tgagtagctg ggattacaag tgtgcgccac cactcctagc taacttttat atttttagta 2340
gagatggtgt ttcaccatgt tgcccaggct ggtcttgaac tcctgacctc aggtgatcca 2400
tctgccttgg tctcccaaag tgcggggatt acaggcatga gccactgcgt ccggccagat 2460
cttcttattt atattgggct aagcatcata tctggcaaaa tgaaagcaag tctttgttat 2520
atgttcagtt tgctactgtc ttccatctcc tggttccttt gactctatgt cttcatcctt 2580
ttccaagaag tctcttgcac atcttccttc ccacaccaca aaagcatctc ttgcacgttt 2640
gggagtttgt gttgtccagg cggatgctgt aggttcaact tcagccaata aaacttcctc 2700
tcacctgagg ctataggtgt ctctgggatc agtaagtcac ttactgagcc tactttctta 2760
ccccagattc tgtgcaccaa atcggagaac aataggctgg ttgtgcaaat agacaatgcc 2820
aaattggctg cagatgactt caggaccaag tgagttgggg ttggacaggg atgggaagga 2880
gccaagtgag agttggggtt ggatggggat gggacaatgc caagttggct gcttctgtct 2940
catgtggcat cggatgtcat tgacccagac cccaacccag tagttatagc agaagagaac 3000
aaagggcatt aagctcaaga tcctgaaaaa ctatccgctc tccagtggtt tagcattaac 3060

<400> 349
ggaaaggaaa ctatgctgta tgccaagccc ccacccacaa ttaatggtat aaaaggactg 60
cagaggaagg agagactcaa acctgcccac atccacctcc agcagcttac ctgcttttcc 120
attacctgtt ccagcaccat gtcttacagt tgttgcctgc ccagcctggg ctgccgcacc 180
agctgctcct cccggccctg cgtgcccccc agctgccacg gctacaccct gcctggggcc 240
tgcaacatcc ccgccaatgt gagcaactgc aactggttct gtgagggctc cttcaatggc 300
agcgagaagg agactatgca gttcctgaac gaccgcctgg ccagctacct ggagaaggtg 360
cgtcagctgg agcgggacaa cgcggagctg gagaaactca tccaggagcg gtcccagcag 420
caggagccct tgctgtgccc cagctaccag tcctacttca agaccattga ggagctccag 480
cagaagattc tgtgtgccaa ggctgagaat gccaggctgg tggtgaacat tgacaatgcc 540
aagctggcct ctgacgactt cagaagcaag taccagacgg agcagtccct gaggctgttg 600
gtggagtcgg acatcaacag catacgcagg atcctggatg agctgaccct ctgcaagtct 660
gacctggagt cccaggtgga gtccctgagg gaggagctga tctgcttgaa gaagaaccat 720
gaggaggagg ttaacaccct gcgctcccag cttggagacc gcctcaacgt ggaggtggac 780
actgccccca ctgtggacct gaaccaggtc ctgaacgaga ccaggagtca gtatgaggct 840
ctggtggaaa ttaaccgcag ggaagtggag caatggttcg ccacgcagac cgaggagctg 900
aacaagcagg tggtatccag ctcagagcag ctgcagtcct gccaggcgga gatcatcgag 960
ctgagacgca cagtcaacgc cctggagatc gagctgcagg cccagcacaa cctgcgagac 1020
tctctggaaa acacgctgac ggagagcgag gcccactaca gctcccagct gtcccaggtg 1080
cagagcctga tcaccaacgt ggagtctcag ctggcagaga tccgctgtga cctggagcgg 1140
cagaaccagg agtaccaggt gctgctggac gtgcgtgccc ggctggagtg tgagatcaac 1200
acgtaccgga gcctcctgga gagtgaggac tgcaagctcc cctgcaaccc atgcgccacc 1260
accaatgcta gtggcaactc ctgtggaccc tgtggcacct ctcaaaaggg ttgctgtaat 1320
tgaaaagctt gtatcctctt tgaagacatc tacaaagcca tttagatcaa ccacaggaag 1380
gatcctcaag tcctgacttt tctggagctc agctgacatc aagaaacctc atcttgcctc 1440
tatgttattt ctagaatgct gaaaagcttt cctgacccaa gcaaagacac acatcatcaa 1500
cttccaatgt ctggacaact ccttcctgtt gagggtcgag cctgtttgtt tctaaagatg 1560
ttcagctccc tgtaatctga gctccagtta ctacttaagg tgtttcctga acgtactact 1620
gcatttcctg ttttcctttt ttctttggca ttctctggaa tgcaaggagg agacttcatt 1680
tacttcccaa taaacttcat ttctctggca taa 1713

<210> 350

<211> 1616

<212> DNA

<213> Homo sapiens

<400> 350
gagaatttag actctgtctt cagccaggca ctccctccct ccctcccagc actatgccct 60
acaacttctg cctgcccagc ctgagctgcc gcaccagctg ctcctcccgg ccctgcgtgc 120
cccccagctg ccacagctgc accctgcccg ggcctgcaa catccccgcc aatgtgagca 180
actgcaactg gttctgcgag ggctccttca atggtagcga gaaggagact atgcagttcc 240
tgaacgaccg cctggccagc tacctggaga aagtgcgtca gctggagcgg gacaacgcgg 300
agctggagaa cctcatccgg gagcggtctc agcagcagga gcccttgctg tgccccagtt 360
accagtccta ttttaagacc attgaggagc tccagcagaa gatcctgtgt accaagtctg 420
agaatgccag gcttgtggtg cagatcgaca acgccaagct ggctgcggat gatttcagaa 480
ccaagtacca gaccgagctg tccctgcggc agctggtgga gtcggacatc aacggtctgc 540
gcaggatcct ggatgagctg accctgtgca agtccgacct ggaggcccag gtggagtccc 600
tgaaggagga gctgctctgc ctcaagagca accatgagca ggaggtcaat accctgcgct 660
gccagcttgg agaccgcctc aatgtggagg tggatgctgc tcccactgtg gacctgaatc 720
gggtgctgaa cgagaccagg agtcagtatg aggccctggt ggaaaccaac cgcagggaag 780
tggagcaatg gttcaccacg cagaccgagg agctgaacaa gcaggtggta tccagctcag 840
agcagctgca gtcctaccag gcggagatca tcgagctgag acgcacagtc aacgccctgg 900
agatcgagct gcaggcccag cacaacctgc gagactctct ggaaaacacg ctgacagaga 960
gtgaggcccg ctacagctcc cagctgtccc aggtgcagag cctgatcacc aacgtggagt 1020
cccagctggc ggagatccgc agtgacctgg agcggcagaa ccaggagtac caggtgctgc 1080
tggatgtgcg tgcccggctg gagtgtgaga tcaacacata ccggagcctg ctggagagcg 1140
aggactgcaa tctgcccagc aatccctgtg ccacgaccaa cgcgtgcagc aagcccatcg 1200
gaccctgtct ctccaatccc tgtacctctt gtgtccctcc tgccccctgc acaccctgtg 1260
ccccacgccc ccgctgtggg ccctgcaatt ccttcgtgcg ctagaaccta gggaatgcca 1320
gaggagcaag gatgcagggc ccaggactcc agagctgtga cctggctctg gttcaacaaa 1380

ctgcaccctg cccggggcct gcaacatccc cgccaatgtg agcaactgca actggttctg 180
tgagggctcc ttcaatggca gtgagaagga gaccatgcag ttcctgaacg accgcctggc 240
cagctacctg gagaaggtgc gtcagctgga gcgggacaac gcggagctgg agaacctcat 300
ccgggagcgg tcacagcagc aggagccctt ggtgtgtgcc agctaccagt cctacttcaa 360
gaccattgag gagctccagc agaagatcct gtgcagcaag tctgagaatg ccaggcttgt 420
ggtgcagatc gacaatgcca agctggcctc agatgacttc aggaccaaat atgagaccga 480
gctgtccctg cggcagctgg tggagtcgga catcaatggc ctgcgcagga tcctggatga 540
gctgaccctg tgcaggtctg acctggaggc ccaggtggag tccctgaagg aggagctgct 600
gtgcctcaag cagaaccatg agcaggaggt taacaccctg cgctgccagc ttggagaccg 660
cctcaacgtg gaggtggacg ctgctcccac tgtggacctg aaccaggtcc tgaatgagac 720
caggagtcag tatgaggccc tggtggaaac caaccgcagg gaagtggagc aatggttcgc 780
cacgcagacc gaggagctga acaagcaggt ggtatccagc tcggagcagc tgcagtccta 840
ccaggcggag atcatcgagc tgagacgcac ggtcaatgcc ctggagatcg agctgcaggc 900
ccagcacaac ctgcgagact ctctggaaaa cacgctgaca gagagcgagg cccgctacag 960
ctcccagctg tcccaggtgc agagactgat caccaacgtg gagtcccagc tggcggagat 1020
ccgcagtgac ctggagcggc agaaccagga gtatcaggtg ctgctggacg tgcgggcgcg 1080
gctggagtgt gagatcaaca cgtaccggag cctgctggag agcgaggact gcaagctccc 1140
ctccaacccc tgcgccacaa ccaatgcatg tgacaagtcc actgggccct gtatctctaa 1200
tccctgtggc ctacgtgctc ggtgtgggcc ttgcaacaca tttgggtact ag 1252

<210> 348

<211> 1621

<212> DNA

<213> Homo sapiens

<400> 348

caggttttga aactgacttc cagagctcca ctgcctccct gcaccatgcc ctacaacttc 60
tgcctgccca gcctgagctg ccgcaccagc tgctcctccc ggccctgtgt gccccccagc 120
tgccacggct acaccctgcc cggggcctgc aacatccctg ccaatgtgag caactgcaac 180
tggttctgcg agggctcctt caatggcagc gagaaggaga ctatgcagtt cctgaacgac 240
cgcctggcca gctacctgga gaaggtgcgt cagctggagc gggacaacgc ggagctggag 300
aacctcatcc gggagcggtc tcagcagcag gagcccttgc tgtgccccag ctaccagtcc 360
tacttcaaga ccattgagga gctccagcag aagatcctgt gcagcaagtc tgagaatgcc 420
aggctggtgg tgcagatcga caatgccaag ctggctgcag atgacttcag aaccaagtac 480
cagacggagc agtccctgcg gcagctggtg gagtccgaca tcaacagcct gcgcaggatt 540
ctggatgagc tgaccctgtg caggtctgac ctggaggccc agatggagtc cctgaaggag 600
gagctgctgt ccctcaagca gaaccatgag caggaagtca acaccttgcg ctgccagctt 660
ggagaccgcc tcaacgtgga ggtggacgct gctcccgctg tggacctgaa ccaggtcctg 720
aacgagacca ggaatcagta tgaggccctg gtggaaacca accgcaggga agtggagcaa 780
tggttcgcca cgcagaccga ggagctgaac aagcaggtgg tatccagctc ggagcagctg 840
cagtcctacc aggcggagat catcgagctg agacgcacag tcaatgccct ggagatcgag 900
ctgcaggccc agcacaacct gcgatactct ctggaaaaca cgctgacaga gagcgaggcc 960
cgctacagct cccagctgtc ccaggtgcag agcctgatca ccaacgtgga gtcccagctg 1020
gcggagatcc gcagtgacct ggagcggcag aaccaggagt atcaggtgct gctggacgtg 1080
cgggcgcggc tggagtgtga gatcaacaca taccggagcc tgctggagag cgaggactgc 1140
aagctgccct ccaacccctg cgccaccacc aatgcatgtg aaaagcccat tggatcctgt 1200
gtcaccaatc cttgtggtcc tcgttcccgc tgtgggcctt gcaacacctt tgggtactag 1260
ataccctggg gccagcagaa gtatagcatg aagacagaac taccatcggt gggccagttc 1320
tgcctctctg acaaccatca gccaccggac cccacccсga ggcatcacca caaatcatgg 1380
tctggaagga gaacaaatgc ccagcgtttg ggtctgactc tgagcctagg gctacttgat 1440
cctcctcacc ccaggtccct ctcctgtagt cagtctgagt tctgatggtc agaggttgga 1500
gctgtgacag tggcatacga ggtgtttttgt tctctctgct gcttctacct ttattgcagt 1560
tccccaaatc gcctaataaa ctttcctctt gcaaagcaga caaaaaaaaa aaaaaaaaaa 1620
a 1621

<210> 349

<211> 1713

<212> DNA

<213> Homo sapiens

<212> DNA

<213> Homo sapiens

<400> 345

```
atgtctggca gttgctcttc taggaaatgc ttctccgtgc cagccacctc tctctgctcc      60
actgaggtga gctgtggagg ccccatctgc ctgcccagtt cctgccagag ccagacatgg     120
cagctggtga cttgtcaaga cagctgtgga tcatccagct gtgggccaca gtgccgtcag     180
ccctcctgtc ctgtgagtag ctgtgcccaa cccctgtgct gtgatcctgt catttgtgag     240
ccttcttgct ccgtgagcag cggctgccaa cccgtgtgct gtgaggccac cacctgtgag     300
ccttcttgct ctgtgagcaa ctgctaccaa cctgtgtgct tcgaggccac catctgtgag     360
ccttcttgct cagtgagcaa ctgctgccaa cctgtgtgct ttgaggccac cgtttgtgag     420
ccttcttgtt ccgtgagcag ctgtgctcaa cctgtgtgct gtgagcctgc tatttgtgag     480
ccttcttgct ccgtgagcag ctgctgccag cctgtaggct ctgaagccac ttcctgccaa     540
ccagtcctct gtgtgcccac ttcctgccag cctgtcctct gcaaatccag ctgctgccag     600
ccagttgtct gtgagcccag ctgctgttca gctgtctgca ccctgcctag ttcctgccaa     660
cctgtggtct gtgagccttc ctgctgtcag ccggtgtgcc cgacacctac ctgctctgtg     720
accagtagct gccaggctgt ctgctgtgac cccagccctt ggtcttctgc atctgccatc     780
tgccgaccaa cttgtcctag gactttctac atacccagtt ccagcaaacg gccttgcagt     840
gctacgattt cctaccgccc ggtctcccgt ccgatctgcc gcccaatctg ctctggactc     900
ctcacctata ggcagccata catgacatcc atctcctacc gtcctgcctg ctatcgccca     960
tgctactcca tcctgcgccg cccagcctgt gtcacttcct actcttgccg cccagtctac    1020
ttccgcccat cttgcactga gtctgactct tgcaaacggg attgcaaaaa atccacttcc    1080
agccaactgg attgtgttga cacaaccccc tgcaaggtgg atgtctcaga agaggctccc    1140
tgccagccca ctgaagccaa acccatcagc ccaaccaccc gtgaggccgc agcagctcag    1200
cctgctgcca gcaagcctgc caactgctaa                                     1230
```

<210> 346

<211> 815

<212> DNA

<213> Homo sapiens

<400> 346

```
gaactccggc tgcgttaccg acctgcgagc tgacgggccc ggaccatggg gtgctgcccg      60
ggggactgct tcacctgctg cacccaggag caaaactgct gtgaagagtg ctgctgtcag     120
ccgggctgct gtggctgctg cggctcctgc tgtggctgtg ggggctctgg ctgcgggggc     180
tctggctgcg gggcagctg ctgcggatcg tcttgctgtg gatctggctg cggaggctgt     240
ggaggctgcg gaggctgcgg gggtggctgc tgtggatcca gttgctgtgg gtccagttgc     300
tgcggctccg ggtgctgtgg gcctgtgtgc tgccagccca cacctatatg cgacacaaaa     360
tgaagacctt tccctccacc actgatgcag tcccaccgaa agcctccatc tgctccaggg     420
ggacagcccc tcgtgtccag aaccttccat acccccaaga cagtgtcttg tctttctgtg     480
atttgtagag gagggcttgt tctccaacac cttctctggt atttcaaggc accgagaaca     540
agagccatac tctgatgaaa cattaaaact cggtcacaac taagtgatcc caagctcaag     600
ggtgaatccc caatacttta cttattcaca tgaagttcaa tgtcttgtac tataagacat     660
cctcttcttc aaggtgtctt tgggactgat cctccgccct ggctttctgc agctttgaga     720
tgcaaaaaag gtccatcttc tttgtgagcc tcttaataaa tttgagcatg ctggcataaa     780
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa                               815
```

<210> 347

<211> 1252

<212> DNA

<213> Homo sapiens

<400> 347

```
ggactctgtc ttcagctgga cactccctcc ctgcaccatg tcttacagtt gtggcctgcc      60
cagcctgagc tgccgcacca gctgctcctc ccggccctgt gtgccccca gctgccacgg     120
```

gctgccagcc ctgctgccgc cccgcctgct gtgagaccac ctgctgcagg accacttgct 420
tccagcccac ctgtgtgtcc agctgctgcc agccttcttg ctgctgatca agtcccaaga 480
gaacaaccat cttcacacaa caaccttctg ctcaactgac ttatcttttg gaggactaat 540
ttaccttact gctgacagca accatgttct cacccaaatt tttatgaatt ctctgcatat 600
ttaaaatctt gtgaatcagc ttgagggagg gcagaatact tcatcctgat tctctttttc 660
ttataccttg tgaatcatgt gccagcttca tctgttctca atttgagtc atggtctcag 720
ctttgactca aaagtcaaga gcttcattct cttcacttaa gaaacttaag ttgctgcaaa 780
tgattaagaa tcttcacaac tatgttttct tttcaatata ctcatgattc ttgtatcctg 840
cttccttctt ttaatgatca ctttgggtta tctccctata accagggatc ttacctatat 900
atttcttaat aaataaattt ggaactatta ttcataaaaa aaaaaaaaaa aaa 953

<210> 343

<211> 990

<212> DNA

<213> Homo sapiens

<400> 343

tctgaacaga agcccaccct ctacccctga caccatgacc cactgttgct ccccttgctg 60
tcagcctaca tgctgcagga ccacctgctg caggaccacc tgctggaagc ccaccactgt 120
gaccacctgc agcagcacac cctgctgcca gccctcctgc tgtgtttcca gctgctgcca 180
gccttgctgc cgcccaactt gctgtcaaaa cacctgctgc cagcccacct gtgtgaccag 240
ctgctgccag ccttcctgct gcagcacacc ctgctgccag cccacctgct gtgggtccag 300
ctgtgaccag agcagctcct gtgcacctgt gtactgcaga agaacctgct actaccccac 360
aactgtctgc ctgcctggtt gcctaaacca gagctgtggc tccaactgct gccagccctg 420
ctgccgccca gcctgctgtg agaccacttg cttccagccc acctgtgtgt ccagctgctg 480
tcagcctttt tgctgctgat caagtcccaa gagaaccacc atcctcacac aacaactttc 540
tgctcaactg acttatcttt tgggggacta atttaatttg ctgctgacag ccaccatgct 600
ctcacccaaa ttttatgaa ttctctacct gtttaaaatc ttgggaatct acttgaggga 660
gggcagaata cttcatcctg attctctttt tccttacact ttgtggatca tgtgccagct 720
tcgtgtgttc tcaatttga gtcatggtct cagctttgac tcaaaagtca agagcttcat 780
tctctgcttc taaggaattt aggtttctgc aactgatcaa taatctttgc aatcatattt 840
ttgttttcaa tatcctcctc atggttcttg tatccttctt tcttcttttc ataactttgg 900
gttatgtttc tgctaccagc agagattctt agctatatgt ttctgaataa actctgaacc 960
atcctcaaaa aaaaaaaaaa aaaaaaaaaa 990

<210> 344

<211> 725

<212> DNA

<213> Homo sapiens

<400> 344

actttggagg ccaggtgtat aaaaggtcca gattgcaagg ggtcatcaga ttctgggaaa 60
ctcacctctg aacagaaacc caccctccac ccctgacacc atgacccact gctgttcccc 120
ttgctgtcag cctacatgct gcaggaccac ctgctgcagg acaacctgct ggaagcccac 180
cactgtgacc acctgcagca gcacatcctg ctgccagccc tcctgctgtg tgtccagctg 240
ctgccagcct tgctgccacc caacttgctg tcaaaacacc tgctgcagga ccacctgctg 300
ccagcccacc tgtgtgacca gctgctgcca gccttcctgc tgcagcacac cctgctaaca 360
gcccacctgc tgtgggtcca gctgctgtgg ccaaatcatc tgtgggtcca gctgctgcca 420
gcccagctcc tgtgcaccca tctactgcag gagaacctgc taccacccca cgagtgtctg 480
cctgcctggt tgcctaaatc agagctgtgg ctccagctgc tgccagccct gctgccgccc 540
agcctgctgt gagaccacct gctgcaggac cacttgcttc cagcccactt ctgtgatcag 600
ctgctgtcag ccttcttgct gctgaccaac tctccagagg accaccatcc tcacacagca 660
accttctggc aaccttctgt cctcctcttg gaggacaaat ttactttcaa actttgctga 720
caacc 725

<210> 345

<211> 1230

<400> 340
tcacctctta acagaagccc accctccatc cctgacacca tgacccactg ttgctcccct 60
tgctgtcagc ctacctgctg caggaccacc tgctggcagc ccaccactgt gaccacctgc 120
agcagcacac cctgctgtca gccctcctgc tgtgtttcca gctgctgcca gccttgctgc 180
cacccaactt gctgtcaaaa cacctgctgt aggaccacct gctgccagcc catctgtgtg 240
accagctgct gccagccttc ctgctgtagc acaccctgct gccagcccac atgctgtggg 300
tccagctgtg gtcagagcag ctcctgtgca cctgtgtact gcagaagaac ctgctaccac 360
cccacaagtg tttgtctgcc tggttgccta aaccagagct gtggctccaa ctgctgccag 420
ccctgctgcc gcccagcctg ctgtgagacc acctgctgca ggaccacttg tttccagccc 480
acctgtgtgt acagctgctg ccagccttct tgctgctaat caactcccaa gagaactacc 540
atcctcacac aacaaccttc agctcaactg acttgtcttt tgagggacta atttactttg 600
ctgctgacag ccaccatgct ctcacccaaa tttttatgaa ttctctacat gtttaaaatc 660
ttgggaatct gcttgaggga gggcagaata cttcatcctc attccctctt tccttacacc 720
ttgtggatca tgtgccagct tcgtctgttc ttaatttgga gtcatgatct cagctttgtc 780
tcaaaaatca agagcttcat tctttgcttc taaggaattt aggtttctgc aactgatcaa 840
tcatctttgc aattatattt tcattttaaa tatccttctc atggttcttg tatccttctt 900
tcttcttttc acgataactt tgggttatgt ctctggtagc agagattctt acctatatgt 960
ttctgaataa actctgaacc atcttcatct caaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1020
<210> 341

<211> 1007

<212> DNA

<213> Homo sapiens

<400> 341
tctgaacaga actccacccct ctacccctga caccatgacc cactgttgtt ccccttgctg 60
tcagcctacg tgctgcagga ccacctgctg gaagcccacc actgtgacca cctgcagcag 120
cacaccctgc tgccagccct cctgctgtgt gtccagctgc tgccagcctt gctgccgccc 180
aacttgctgt caaaacacct gctgccagcc catctgtgtg accagctgct gccagccttc 240
ctgctgcagc acaccctgct gtcagcccac ctgctgtggc caaaccagct gtgggtccag 300
ctgtggtcag agcagctcct gtgcacctgt gtactgcaga agaacctgct accaccccac 360
gactgtctgc ctgcctggtt gcctaaacca gagctgtggc tccagctgct gccagccctg 420
ctgccgccca gcctgctgtg agaccacctg ctgcaggaca acttgcttcc agcccacctg 480
tgtgtacagc tgctgccagc cttcttgctg ctgatcaagt cccaagagaa ccaccatact 540
cacacaacaa atttctgctc aactgactca tcttttgggg gactaattta atttgctgct 600
gacagccacc atgctttcac ccaaattttt atgaattctc tgcatgttta aaatcttgtg 660
aatcagcctg agggagggca gaatacttca tcttgattct ctttttcctt acaccttgtg 720
gatcatgtgc cagcttcgtc tgttctcaat ttggaatcat gatctcaact ttgactcaaa 780
agtcaagagg ttcattctct gcttctaagg aatataggtt tctgcaaccg accaataatt 840
tttgcaatca catttttgtt ttcaatatcc tcctcatggt tcttgtattc ttctttattc 900
ttttcatgat aactttgagt tatgtccttg gtaacagaga ttcttaccta tatatttctg 960
aataaactct taaccatcct catctcaaaa aaaaaaaaaa aaaaaaa 1007
<210> 342

<211> 953

<212> DNA

<213> Homo sapiens

<400> 342
ccatgaccca ctgttgctcc ccttgctgtc agcctacctg ctgcaggacc acctgctgca 60
ggaccacctg ctggaagccc accactgtga ccacctgcag cagcacaccc tgctgccagc 120
cctcctgctg tgtgtctagc tgctgccagc cttgctgccg cccagcttgc tgtcaaaaca 180
cctgctgcag gaccacctgc tgccagccca cctgtctgtc cagctgctgt ggccaaacca 240
gctgtgggtc cagctgtggc cagagcagct cctgtgcacc tgtgtactgc agaagaacct 300
gctactaccc gacgactgtc tgcctgcctg gttgcctcaa ccagagctgt ggatccagct 360

ttattaacca gcattcttga tatgatccac ctgtgaactg aatcatgcaa ggccaattgg 540
acaacctcag ttccaaccaa ttcttggatt gagtttggcc tccaaatatg ctcaccaaac 600
actatgttgc taccctctac caaatgaata caagtttgaa ttttctctga aatatgtcaa 660
ccccatgttc cctgaattga aatatttgct ctctaccata atttatcaca tggagctatt 720
cctctatctt aaataaattt taatttttga ggcatgcaaa taaaaaaaaa aaaaaaaaaa 780
aa 782

<210> 338

<211> 754

<212> DNA

<213> Homo sapiens

<400> 338
ctgtcttgga aacctaccta gaacctccac cctctgacac catggttaac tcttgttgtg 60
gctctgtctg ctctgaccag ggctgtgatc aaggcctctg ccaagagacc tgctgccgcc 120
ccagctgctg ccagaccacc tgttgctgcc ccagctgtgt tgtatccagc tgctgccgcc 180
catcctgctc tcagactacc tgctgccaga ccacttgctg tcgccccagc tgctgccgcc 240
cagtctgttg tcagaccacc tgccgcccca gctgtggtgt gtccagctgc tgccgtccac 300
tctgttgtca gaccacctgc cgccccagct gtggtgtgtc cagctgctgc cgtccactct 360
gctgtcagac cacctgctgc cgtacaactt gctgccgccc cagctgctgt ggatcctctt 420
gttgaacctc atattggact atcaaccatg agccagtcac catcccatga tatgaaagtg 480
tcatttatat caatttgtcc atgtgttaaa tggccatcag ctttattatc ctgttattcc 540
actaagtaac tggtgaaagg gcatccatct acatttaata ccactcattt ttcccatgga 600
tctcttccca gcattcagac ctcagatgca tgatgcagtt agactaagcc tctaagtctc 660
tgactgttat gcttgatttg accttcaaaa ttacatatta actgttcttc aataaatatt 720
atcttaatat cagtaaaaaa aaaaaaaaaa aaaa 754

<210> 339

<211> 1042

<212> DNA

<213> Homo sapiens

<400> 339
tgacaccatg acccactgtt gctccccttg ctgtcagcct acctgctgca ggaccacctg 60
ctgcaggacc acctgctgga agcccaccac tgtgaccacc tgcagcagca caccctgctg 120
ccagcccgcc tgctgtgtgt ccagctgctg ccagccttgc tgccgcccaa cttgctgtca 180
aaacacctgc tgtaggacca cctgctgcca gcccacctgt gtgaccagct gctgccagcc 240
ttcctgctgc agcacaccct gctgccagcc cacctgctgt gggtccagct gctgtggcca 300
aaccagctgt gggtccagct gtggccagag cagctcctgt gcacctgtgt actgcagaag 360
aacctgctac taccccacga ctgtctgcct gcctggttgc ctaaaccaga gctgtggctc 420
caactgctgc cagccctgct gccgcccagc ctgctgtgag accacctgct gcaggaccac 480
ttgcttccag cccacctgtg tgtccagctg ctgccagcct tcttgctgct gatcacgttc 540
caagagaacc accatcctca cacaacaaat ttctgctcaa ctgactcatc ttttgggga 600
ctaatttaat ttgctgctga cagccaccat gctctcaccc aaatttttat gaattctcta 660
catgtttaaa atcttggaaa tctgcttgag ggagggcaga atacttcatc ctgattctct 720
ttttccttac accttgtgga tcatgtgcca gcttcatctg ttctcaagtt tgagtcatgg 780
tctcagcttt gactctaaag tcaagagctt cattccctgc ttctaaggaa tttaggtttc 840
tgcaactgat cgatgatctt tgcaatcttt tttttgtttt caatatcctc ctcatcgttc 900
ttgtatcctt ctttcttctt ttcatgataa atttgtgttg tgtccctggt agcagaaatc 960
cttacctata tgtttctgaa taaattctga accatcctca tctcaaaaaa aaaaaaaaaa 1020
aaaaaaaaaa aaaaaaaaaa aa 1042

<210> 340

<211> 1020

<212> DNA

<213> Homo sapiens

<400> 335
ctccttcctg gaaacccacc cagaacctcc accctctgac accatggtca actcctgttg 60
tggctctgtg tgctctgacc agggctgtgg cctagagaac tgctgccgtc ccagctactg 120
ccagaccacc tgctgcagga ccacctgctg ccgccccagc tgctgccgcc ccagctgctg 180
caggccccag tgctgccagt ctgtgtgctg ccagcccacc tgctgctgcc ccagctactg 240
tgtgtccagc tgctgcagac cccagtgctg ccagaccacc cgctgcagaa ccacctgctg 300
ccgccccagc tgctgtgtgt ccaggtgcta caggccccat tgtggccagt ctctatgctg 360
ctagcccatc tgctggcaaa ccacctgcta caggaccacc tgctgccacc ccagctattg 420
catttccagc tgctgccagc cttcctgcag tatctctagc agcagtagct cctcctgctg 480
tggcttcagc tgctgcaggc tctcctgctg catctccagt tgctgccgcc ccaactgctg 540
ccagaccatg tgctgccgcc caacctgctc tagtgcttct tgctgctgag gctgtcatct 600
ggactcacca gattctcatc aaccagcatt cttgatgtag ctcatctatg agctgagtta 660
tgggaagcta gttggaaaac ttcagttcca accaattctt agattgaatc tggcctccaa 720
atatatgctc cccccacatt ttacctctct accaaatgaa cataagttga atttgctctg 780
aaatctgtca actatcttaa ttgaaatatt tgctctctgc cataatttct catatggagc 840
tattccattt taaacaaata tttatctaaa taaatcttaa ataaattttc aggcatagaa 900
ataaaaaaaa aaaaaaaaaa aaaa 924

<210> 336

<211> 1099

<212> DNA

<213> Homo sapiens

<400> 336
cttcctggaa acccacccag aacctccacc ctctgacacc atggtcaact cctgttgtgg 60
ctctgtgtgc tctgaccagg gctgtggcct agagaactgc tgccgtccca gctactgcca 120
gaccacctgc tgcaggacca cctgctgccg ccccagctgc tgtgtgtcca gctgctgcag 180
accccagtgc tgccagacca cctgctgcag gaccacctgc tgccacccca gctgctgtgt 240
gtccagctgc tgcagacccc agtgctgcca gtctgtgtgc tgccagccca cctgctgcag 300
accccaatgc tgccagacta cctgctgtag gaccacctgc tgccgcccca gctgctgcag 360
gccccagtgc tgccagtctg tgtgctgcca gcccacctgc tgctgcccca gctactgtgt 420
gtccagctgc tgcagacccc agtgctgcca gaccacctgc tgcagaacca cctgctgccg 480
ccccagctgc tgtgtgtcca ggtgctacag gccccattgt ggccagtctc tatgctgcta 540
gcccatctgc tggcaaacca cctgctacag gaccacctgc tgccacccca gctattgcat 600
ttccagctgc tgccagcctt cctgcagtat ctctagcagc agtagctcct cctgctgtgg 660
cttcagctgc tgcaggctct cctgctgcat ctccagttgc tgccgcccca actgctgcca 720
gaccatgtgc tgccgcccaa cctgctctag tgcttcttgc tgctgaggct gtcatctgga 780
ctcaccagat tctcatcaac cagcattctt gatgtagctc atctatgagc tgagttatgg 840
gaagctagtt ggaaaacttc agttccaacc aattcttaga ttgaatctgg cctccaaata 900
tatgctcccc ccacatttta cctctctacc aaatgaacat aagttgtaat ttgctctgaa 960
atctgtcaac tatcttaatt gaaatatttg ctctctgcca taatttctca tatggagcta 1020
ttccatttta aacaaatatt tatctaaata aatcttaaat aaattttcag gcatagaaaa 1080
aaaaaaaaaa aaaaaaaaa 1099

<210> 337

<211> 782

<212> DNA

<213> Homo sapiens

<400> 337
ccaccttcct ggaaatccac ccagaacctc caccctctga caccatggtc aactcctgtt 60
gtggctctgt gtgctctgac cagggctgtg gcctagagaa ctgctgccgt cccagctgct 120
gccagaccac ctgctgcagg accacctgct gccgccccag ctgctgtgtg tccagctgct 180
gcagaccgca gtgctgccag tctgtgtgct gccagcccac ctgctgcagc cccagctgct 240
gccagaccac ttgctgcagg accacctgct gccgtcccag ctgctgtgtg tccagctgct 300
tcagacccca gtgctgccag tctgtgtgct gccagcccac ctgctgccgc cccagctgtg 360
gccagaccac ctgctgcagg accacctgct accgccccag ctgctgtgtg tccacctgct 420
gccgcccaac ctgctctagt ggctcttgct gctgatgccc tcacctatac tcacctgcct 480

<210> 333

<211> 1077

<212> DNA

<213> Homo sapiens

<400> 333
```
ctcactctcc tggaaaccca ccgagaacct ccaccctctg acaccatggt caactcctgt      60
tgtggctctg tgtgctctga ccagggctgt ggcctggaga actgctgccg ccccagctgc     120
tgccagacca cctgctgcag gaccacctgc tgccgcccca gctgctgtgt gtccagctgc     180
tgcaggcccc agtgctgcca gtctgtgtgc tgtcagccca cctgctgccg ccccagctgc     240
tgtcagacca cctgctgtag gaccacctgc tgccgcccca gctgctgtgt gtccagctgc     300
tgcagacccc agtgctgcca gtctgtgtgc tgccagccca cctgctgccg ccccagctgc     360
tgtcagacca cctgctgcag gaccacttgc tgccgcccca gctgctgtgt gtccagctgc     420
tgcagacccc agtgctgcca gtctgtgtgc tgccagccca cctgctgccg ccccagctgc     480
tgcatctcca gcagctgctg cccctcttgc tgtgaatcca gctgctgccg cccctgctgc     540
tgcctgcgtc cagtctgtgg ccgagtctcc tgccacacca cttgctatcg cccaacctgt     600
gtcatctcca cctgcccccg ccccttgtgc tgtgcctcct cttgctgcta aatctctgct     660
gtgaacacac cacttcctta ttacgtcctt tcctacagat gaaggctctc attgcaaaca     720
tgcggactgt tcaagagaat tgatctgggt cccataagca aacctcatcc ttagaaattc     780
tgtatttgca ttctaccttt tgtccaaact cccttccttc caaaggaatt cattgacaat     840
ctcctaataa attgacaaat tgtcctccaa catcctccca cctctttgac ttcaggacat     900
ttattcatca tgcctaagga atttgaagat tgcctccatc atttgtaggg ccacagatct     960
taaagcctcc aaccttgaag tccagtgaag tctctctctt aaagtctttt gcaaacattt    1020
ttgtaccttg ttatttccat gtaccaaaat aaacctctat tctattggca ctgaaaa       1077
```
<210> 334

<211> 942

<212> DNA

<213> Homo sapiens

<400> 334
```
tagaaatcca cccagaacct ccaccctctg acaccatggt cagctcctgt tgtggctctg      60
tcagctctga gcagagctgt ggcctggaga actgctgccg ccccagctgc tgccagacca     120
cctgctgcag gaccacctgc tgccgcccca gctgctgcaa gccccagtgc tgccagtctg     180
tgtgctacca gcccacctgc tgccacccta gctgctgcat ctccagctgc tgccacccct     240
attgctgtga atccagctgc tgccgcccct gctgctgccg ccccagctgc tgccagacca     300
cctgctgcag gaccacctgc tgcaggacca cctgctgctg ccccagctgc tgtgtgtcca     360
gctgctgcag accccagtgc tgccagtctg tgtgctgcca gcccacttgc tgccgtccca     420
gctgctgcat ctccagctgc tgccacccct cttgctgtga atccagctgc tgccgcccct     480
gctgctgcgt gcgtccagtc tgtggccgag tctcctgcca caccacttgc tatcgcccaa     540
cctgtgtcat ctccacctgc ccccgcccct tgtgctgtgc ctcctcttgc tgctaatgtc     600
tccttgtgat atttgtcata ctatgaatgt cttcattagt catttaaaat gcactgtagc     660
cagccagtca ctggaaaaat gaacacttcc ctgccagttt gtctcatgtg gcattcagag     720
tggacattca gctcttctag gaaatgacag acaatcacat tcattaaaat atgttatgcc     780
aggccccaat gcagttattt ttagatgagc agtgtcttca ttcgaagggg acactaactg     840
tgatgatctc atataatatt gttttcatgt attaataaac agccacttcc ctaaaaaaaa     900
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa                        942
```
<210> 335

<211> 924

<212> DNA

<213> Homo sapiens

```
caccagtcta aggggcagga aacctgggct ctcaaacaag ctctgccact agtttgccat   6420
gagacttggg aaaggcattt aacctctctg gcctgttgtt ctagatcagt gatcacacat   6480
tcagacatca atggggttag taaatgagcg aagcaagcaa gctgttccca aggcacttgg   6540
gtatgatagg gactatggta aactggaaag cacaaaaggc agggctaggc agatccaact   6600
gaggcaggcc atgcagaaat gcaggtccag tgtggccaga tcctctgatc ttttgaaaga   6660
agccaagaat ttcaattatt gtgttaaatt tccaaaattg aagacactgt gcaaacatgc   6720
ctgggagcac taatttgcaa cccctggcct aacaaataga caaggcctct ttttcttctg   6780
ctagtcagtg gtcagtggtt acctggcctc tgaccctgaa gccagaacag agatccttaa   6840
taaccctccc ttgcctccca gaactccaag ctggaggccg cggtggccca gtctgagcag   6900
cagggtgagg cggccctcag tgatgcccgc tgcaagctgg ccgagctgga gggtgccctg   6960
cagaaggcca agcaggacat ggcctgcttg atcagggagt accaggaggt gatgaactcc   7020
aagctggcct ggactttgag atcgccacct acaggcgcct gctggagggc gaggagcaga   7080
ggtgaggacc acagctgtgg ggagggtgtc ccctcagttt cccccaggca gtgctgcagt   7140
gcttcaacac ccaggagcat gtctgcatat tatttacttg gaatgtgcat caacgctaat   7200
tggaggagag aatttgcagg ttaaacagat ttcaccagct ttgctcccca caactaattt   7260
ctgtctcttt gctttcccta taaggttgtg tgagggcgtt ggtgctgtaa atgtctgtaa   7320
gtaatccatt ttgtgggtcc tgaaataaca gcttggggtt ctgggcaggc aaactttttc   7380
ttagagcttc agatagtaaa aactttaggc tcgtgggcca tgtagtctct gcctcaactg   7440
ctccactccg ttgtcatgta aagcagccac agccaaaggc aatacttgaa gaagcggtat   7500
gactgggttc caagaaaact tgattgacaa acaggaagcc tgcatgtggc cgtagtttgc   7560
agactctaaa aggaaagatg gagtggggca gtgagtgctt atagtaactt aattagctac   7620
agtcatgtac cacataagga cttttaggtc aaggatggat ggaatatagg acagtagtcc   7680
cattagatta taccatattt ttactctact ttcctaggtt tagctatgtt cagatacaca   7740
aatacctacc atggtgttac agttgccttt agtcttcagt acagtcacat gctgtacaga   7800
tttgtagctt agaagcaaca ggctatgcta tagggtaaac ttgtcaaacc cctgaccaga   7860
aggctgcata cagcccagga cggctttgaa tgtggcccag cacaaatttg taaacattct   7920
taaaacttac gagattttcg tgtgtgtgat tttttaaagc tcatcagcta ccgttagtat   7980
tagtgtattt tatgtgtggc ccaagacaat tcttcttctt ccagtgtggc ccagggaagc   8040
caaaagattg gacacccttg gggcctgggt gtggtggctt acacctgtaa tcccagcact   8100
ttgggaggcc aaggcaggtg gatcacttga ggtcaggagt ttgagatcag cctggccaac   8160
atggtgaaac cccgtctcta ctaaaaataa aaaaattagc cgggcaaggt gatggacacc   8220
tgtaatccca gctactaggg aggctgaggc aggagaatca cttaaaccca ggaggcagag   8280
gttgcagtga gccgagatcg cgccactgca ctccagtctg ggcgatagag agagactccg   8340
tctcaaaaat aaaaataaaa aataaaaaag gttggccacc cttgccacag agcccaggtg   8400
tgtagtaggc tacaacatgt aggtttgcgt atgtgcactc catgatgttc gcacaatgac   8460
aaaatcattt aatgctgtat ttctcagaac atgtctctgt tgtgaggcga agcatggcta   8520
tgtctgccta tagagaggac ctgggctaac agtgtatcaa ggacgcctga ggctaagact   8580
gcaggaagga gtgcccatgt gagggtttga ggggcgcgca tagaaacaga ggggagaagc   8640
tctgagcaga aggggtctct ggccgccagg agactgggcc ctgtcctccc cctctggcag   8700
gtgtcagcag ctcccgcggc ggggtcttat gcggggacct gtgcgtgtcg ggctcccggc   8760
cggtgacggg cagcgtctgc agtgcccccct gcagcgggaa cgtggcggtg agcaccggcc   8820
tgtgtgcgcc ctgcggccag agcagcggca gcggccgctc cgtgcgcttc gcctgagcgc   8880
cccactttgt accgacgccg cctcaccagc agcagcctct gcctccagag gcagcccgga   8940
gctcccacgg tccccagacg tggggaagga gggcttcggg gcccagcctc cccatcggcc   9000
cccacgccgc gccttcctgt ctcgcatccc cgcccgctag tcccttcact gcttccgaga   9060
gcacagggct cctagcccgc agctctccca acccccgaat ctaaaggact gtcctcgcag   9120
ccctaggtaa agaccccggt ccggaggcct ggatgcctgt ggcttctgag ggacagatgg   9180
gcctggagcc ccttggtgtg gacttgctgt gccttctaat ccaatgtctt ctaatacacg   9240
tgcaggacag gccacctctc ctgcctggct gcctggctcc ctccctctcc cctgtgcttg   9300
tctcagcatt tccagtaaag ctcctgatca tatgcattgt gccctgagag tctctccttc   9360
ctcttgccca cttggtcttc ttagtcggga gggagagggc ctggatctcc tgacaccttg   9420
ggtgggaacc ccatcactgt ggctgctggt ccactttttc cactctccac cccagccccc   9480
acctcagcat ttgaaagttc cctgaccata gggtcaaata tacagacttc ctcccagaag   9540
ccaacacatc ccacatagct cccccatcca caccaccacc atcaccccctg gggtcccaga   9600
cacagagacc tccctgtggc tccacagtac ccatcatatg tgctgctgcc cacatggcca   9660
catttcccca tgcataggcc ctgccacaga acacctgcca cactcacata gcacacagac   9720
c                                                                   9721
```

<210> 366

<211> 4721

<212> DNA

<213> Homo sapiens

<400> 366

```
tgctagcagt gagggaaagg tcagagtcct cttcttcctc tggcttcatg tcctccagca     60
gagagtcgcc atgacttgta gattgtcttg cctaggctcc cgagcggagt ctgcaaatcc    120
aggtcctgtg cagccactca gagatatggt tatgacccca ccctcagctg catcttgccc    180
tccgggccct ggtctggtgg ctgctgcatc acggcaagcc cctgctacca tggcctcact    240
gggggctttg gcactggttc ctgtggacac agattcagct accactctgg cagcacatgc    300
gggcccagca ctgcctgtaa caccattgtg ttgatcaatg agagcctcct catgcccctc    360
aacctggaga ttgcacagtg cataaagcag gaggaggaga agatcaagtg cctcaacagc    420
aggtttactg ccttcctcca caaggtgggt gtcctggatc acacccttcc tgaaccccca    480
ccatgtgcac aaccaggact gggcactgag gaaagaatca gaggcaggca agaccctgcg    540
tgcctgaggt cccagtctga tgggagaggc acacagacag acagacagac agacagacag    600
acagacagac agaggcaccc aaggcccaga aatactctga cagttcagat acggagagct    660
tgtggaggaa agatgccaaa tttcaaaatt tcaatgaatt tagtttaaag atcttagttg    720
gattttatat gcaattttag aattgggcag cccttagaac tgatcagagg aactctggct    780
gcactgcggt cccacaatac ttatggacag aaaacagaaa gtgacacatg gaaaacagaa    840
gtaacagctg gattgtttgc agctggtgtt tgtgaacatt tggctgcctg tgattgactg    900
aagctcagtt actgtgattg gctgagactc agctatttgt tacaaaggaa aattctgaag    960
ttaggtttgt aagggataat tggaaggaca gccgagaaag gaataaggcc aatagaccca   1020
agttcaggca agctgattta ctgtcagtcc tgccgggcta cctcctgaca aaagcagagg   1080
aggcagccct gcttataggc tattgcaagg ttttataggg catgtattcc taattccaac   1140
taggaatgtt ggaattagca gtttctggcc aaggatctga ggtacagtta ttataggtta   1200
aaagagtgct ggccaggcgc ggtggctcac gcctgtaatc tcagcacttt gggaggctga   1260
gcgggggtag atcacgaggt caggagtttg agaccagcct gaccaacatg gtgatacccc   1320
gtctctacta aaaatacaaa aattagcttg gcgtggtggc acgcacctgt aatcccagct   1380
actccagagg ctgaggcagg agaattgctt gaatctggga ggcagaggtt gcagtgagct   1440
gaggttgcgc cactgtactc cagcctgggt gacaaagcaa gattctgtct caggaaaaaa   1500
aaagtgctat agtaccctgg cacataagtc aatttgcaag aaaatctgtt tgggattttg   1560
gtgcaaactt gccttgcaat atactcatta gtcttggttc ttgaatcaca tgactgggtt   1620
aaagtaaaac aggtcatgag aacaaatgag tctggaagtc taattccgta atagggacca   1680
ctgctataat acatcataga ttgatccctt atgaaatagc tactggaagg cctatgccct   1740
aataatagaa ctcatttatt ttctactccc ctaaactctg acataactaa atgctgcaag   1800
gttttaatgc attatgccaa agtatatttt caccaggtaa aggaagcttt tcaagatcta   1860
tcaactgagg acaatcaaac ccttcacaat ctagaatgca gagatcaggt catctggaga   1920
cgatatctga gaaagactgc cattgagccc cattggaggg ggccatatca agttcttctc   1980
atcacccact cttcagcaaa acttcagggc cttggacctt cggtctacat ctcgcaactc   2040
aaaagggtcc attcagactc ttagaactgc acatccattg gagaccttaa gataaagcca   2100
agcagagaag tttcttccca gaaacagaca gtattctaga tgtggacagc tttcccaaga   2160
ccatggatca agacttctct gccgtcatgc aactcttacc tctcttaatt ttttccttgc   2220
ttacgccttc ctcattcact tggcaggata atgctgtaat tcgaatttca taatcagtag   2280
cttctgagag taattttttt ttctttttga gacggagtct ctattgccca ggttggggtg   2340
cagtggtgtg atcttggctc actgcaaacc tctgtctctt gggttcaagc aattctcatg   2400
cctcagcctc ccaagtagct gggattactg gcatgcacca acacactggc taatttttgt   2460
attttttttt tttttttttt ttggtagaga tgggatttca ccatgttggc catgctggtc   2520
acaaactcct cactgcaggt gatctgcctg cctcagcctc ccaaagtgct aagattacag   2580
gagtaggccg ccatgcccaa ccttctgaga gtaacttgac ggagtgttgt atctgtcatg   2640
tcaaactcaa atctttacat tacctacaaa gtccacctat tggccaactt cagcaacatc   2700
cctaatgcaa ctatttgttc aaattatagc tatgttctca acttaagtca ctgcccccac   2760
ctcttgcagg agatccatgt tctccaatcc cacatctcag acacctctgt ggttgtcaag   2820
ctggacaaca gccaggacct gaacatggac tgcatcattg ccgagatcaa ggcacagtat   2880
gacgacattg tcacctgcag ctgggccaag gctgagtctt ggtaccatgg ctgcctgcca   2940
gatgtggcac ctgcctgcta gatgtggcag ttgggagggc aggatgtgag atacctatta   3000
gatagactta ttttgcctgg ggattctgga tagtaggcca atagctctca ggttggggtg   3060
gcagggcagg actgccatgt gtggttgcac aggctgagca ctgcacaacc tgcacaatca   3120
tccctaaagg aagggatatc ccacactgag cctcaggagc atctctgctt cccccatttc   3180
ggggagatga aggccacggt gatctggcac ggggagactg tgggctgcac caaggaggag   3240
atcaaggagc tgacccacat gatccagagg ctgatggcca aggtggagaa tgccaagtgc   3300
caggtatggg gcatctgtgc ccaaggccag agagacttat ggcctaacct ttgtcacaca   3360
gcctatgtgt gccctacgtg gatctcagca ttcattctcc agtccctttg tccatgtaga   3420
gtccctggtt gtggtcagtc agggagtgcc aggatgatga aggcaagagg cctgtttctg   3480
aggtgctccc agctgaatgg cagactggac atattcaggt aatggacaga gagaccgtaa   3540
cctatcttgt tctcttctgg tcccccagaa ctccaagctg gaggtggagg tgacccagtc   3600
tgagcagcag ggtgaggtgg cccttagcga tgaccgctgc aagctggtcg agctggaggg   3660
caccatgcag aaggccaggc agtaccagga ggtgatgaac tccaagctgg gcctcaatgt   3720
gaaggctgcc tcttcttgcc aacttctaga tggtgaggag cagaggtggg tgcagcctgc   3780
agagctgggc ccagagctca tctgcataga cctgccccac tttttagcac catcacttac   3840
```

ccagccctca ctactacttc tgttttcatc ttggaaccct tttaggacct ctactggtat 3900
tgagcacagt ccaaggagcg gagtcttagt tgagtccaca ggccgatttc caataggctc 3960
cttgcatgat ctattaggga caatggtttg tctcagagaa cttgttctct gagataaaca 4020
atgccctaat ttataacact taccgatttc agagttgtat atacttctgt catggctgat 4080
ttcaggctgt tgctgtgact tcatgtagaa ttggctagaa atgtgcacaa taggctcttg 4140
tgagctggtg ggagctggtt caagaacaac attggttggg gctaggatgg aatggtgtct 4200
cacactgaaa gacaaatagg aagatatttg tctcttctct tagatgaaaa acctgaggct 4260
tagggagact gtgggggttg gggagtccgc ctggtgtgga gttattggga cagttggaga 4320
gttcttttgt gtcctgctgc attttcacag gctccacttc ctggggggat ctgggagagc 4380
ctggcacttg tgttgttttg ggacagggtt gggacgatgc tgggcctttg agggtgctag 4440
acccaggtct tccctcctct cactttctgg gaagggcttg ggtgacctct tggttgactg 4500
agttcagagc tggtgaccac atggtctctg gtgtcattgt caggaattaa tgcccagtgc 4560
taaagctgac cttttctctc ctgcaccagg ctgtgtgatg gcatcagtgc tgtgaatatc 4620
tgtgagtaat tctggccctg aagggagact ctgaggtttg ccagtgggac gaattttta 4680
gctctgaggg tggaaatacc aaggaagtgg ttccattcat t 4721

<210> 367

<211> 1925

<212> DNA

<213> Homo sapiens

<400> 367
actccaggtc ccctatcctg tcctctgcaa cccaaacgtc caggaggatc atgacctgcg 60
gatcaggatt tggtgggcgc gccttcagct gcatctcggc ctgcgggccg cgccccggcc 120
gctgctgcat caccgccgcc ccctaccgtg gcatctcctg ctaccgcggc ctcaccgggg 180
gcttcggcag ccacagcgtg tgcggaggct ttcgggccgg ctcctgcgga cgcagcttcg 240
gctaccgctc cgggggcgtg tgcgggccca gtcccccatg catcaccacc gtgtcggtca 300
acgagagcct cctcacgccc ctcaacctgg agatcgaccc caacgcgcag tgcgtgaagc 360
aggaggagaa ggagcagatc aagtccctca acagcaggtt cgcggccttc atcgacaagg 420
tgcgcttcct ggagcagcag aacaaactgc tggagacaaa gctgcagttc taccagaacc 480
gcgagtgttg ccagagcaac ctggagcccc tgtttgaggg ctacatcgag actctgcggc 540
gggaggccga gtgcgtggag gccgacagcg ggaggctggc ctcagagctt aaccacgtgc 600
aggaggtgct ggagggctac aagaagaagt atgaggagga ggtttctctg agagcaacag 660
ctgagaacga gtttgtggct ctgaagaagg atgtggactg cgcctacctc cgcaagtcag 720
acctggaggc caacgtggag gccctgatcc aggagatcga cttcctgagg cggctgtatg 780
aggaggagat ccgcattctc cagtcgcaca tctcagacac ctccgtggtt gtcaagctgg 840
acaacagccg ggacctgaac atggactgca tcattgccga gattaaggca cagtatgacg 900
acattgtcac ccgcagccgg gccgaggccg agtcctggta ccgcagcaag tgtgaggaga 960
tgaaggccac ggtgatcagg cacggggaga ccctgcgccg caccaaggag gagatcaatg 1020
agctgaaccg catgatccaa aggctgacgg ccgaggtgga gaatgccaag tgccagaact 1080
ccaagctgga ggccgcggtg gctcagtctg agcagcaggg tgaggcagcc ctcagtgatg 1140
cccgctgcaa gctggccgag ctggagggcg ccctgcagaa ggccaagcag gacatggcct 1200
gcctgatcag ggagtaccag gaggtgatga actccaagct gggcctggac atcgagatcg 1260
ccacctacag gcgcctgctg gagggcgagg agcagaggct atgtgaaggc attggggctg 1320
tgaatgtctg tgtcagcagc tcccggggcg gggtcgtgtg cggggacctc tgcgtgtcag 1380
gctcccggcc agtgactggc agtgtctgca gcgctccgtg caacgggaac gtggcggtga 1440
gcaccggcct gtgtgcgccc tgcggccaat tgaacaccac ctgcggaggg ggttcctgcg 1500
gcgtgggctc ctgtggtatc agctccctgg gtgtggggtc ttgcggcagc agctgccgga 1560
aatgttaggc accccaactc aagtcccagg ccccaggcat ctttgcctgc cctgccttgc 1620
ttggcccagt cagtcaggcg cctggagaag tgctcagcta cttctcctgc actttgaaag 1680
acccctccca ctcctggcct cacatttctc tgtgtgatcc cccacttctg ggctctgcca 1740
ccccacagtg ggaaaggcca ccctagaaag aagtccgctg gcacccatag gaagggcct 1800
caggagcagg aagggccagg accagaacct tgcccacggc aactgccttc ctgcctctcc 1860
ccttcctcct ctgctcttga tctgtgtttc aataaattaa tgtagccaaa aaaaaaaaa 1920
aaaaa 1925

<210> 368

<211> 2108

<212> DNA

<213> Homo sapiens

<400> 368
tccattcgga cgtctccatc ctcagaacct cctctcttcc ccaaaaagca ccatgacttg 60
tggatcttac tgtggtggcc gcgccttcag ctgcatctcg gcctgcgggc cgcgccccgg 120
ccgctgctgc atcaccgccg cccccctaccg tggcatctcc tgctaccgcg gcctcaccgg 180
gggcttcggc agccacagcg tgtgcggagg ctttcgggcc ggctcctgcg gacgcagctt 240
cggctaccgc tccgggggcg tgtgcgggcc cagtccccca tgcatcacca ccgtgtcggt 300
caacgagagc ctcctcacgc ccctcaacct ggagatcgac cccaacgcgc agtgcgtgaa 360
gcaggaggag aaggagcaga tcaagtccct caacagcagg ttcgcggcct tcatcgacaa 420
ggtgcgcttc ctggagcagc agaacaaact gctggagaca aagctgcagt tctaccagaa 480
ccgcgagtgt tgccagagca acctggagcc cctgtttgag ggctacatcg agactctgcg 540
gcgggaggcc gagtgcgtgg aggccgacag cgggaggctg gcctcagagc ttaaccacgt 600
gcaggaggtg ctggagggct acaagaagaa gtatgaggag gaggtttctc tgagagcaac 660
agctgagaac gagtttgtgg ctctgaagaa ggatgtggac tgcgcctacc tccgcaaatc 720
agacctggag gccaatgtgg aggccctgat ccaggagatc gacttcctga ggcggctgta 780
tgaggaggag atccgcgttc tccagtccca catctcagac acctccgtgg ttgtcaagct 840
ggacaacagc cgggacctga acatggactg catcattgcc gagatcaagg cacagtacga 900
tgacattgtc acccgtagcc gggctgaggc cgagtcctgg taccgcagca agtgtgagga 960
gatgaaggcc acggtgatca ggcacgggga gaccctgcgc cgcaccaagg aggagatcaa 1020
cgagctgaac cgcatgatcc agaggctgac ggctgaggtg gagaatgcca agtgccagaa 1080
ttccaagctg gaggctgcgg tggctcagtc tgagcagcag ggtgaggcgg ccctcagcga 1140
tgcccgctgc aagttggccg agctggaggg tgccctgcag aaggccaagc aggacatggc 1200
ctgcctgatc agggagtacc aggaggtgat gaactccaag ctgggcctgg acatcgagat 1260
cgccacctac aggcgcctgc tggagggcga ggagcagagg ctgtgcgaag gcgtcggctc 1320
ggtgaatgtc tgcgtcagca gctcccgcgg tggcgttgtc tgtggcgatc tctgcgcctc 1380
cactactgcc cctgttgtct ccaccagagt cagtagcgtc cccagcaaca gcaacgtggt 1440
ggtgggcact actaacgcct gcgccccctc cgcccgggtt ggcgtctgcg gcggcagctg 1500
taagaggtgc taaggaggct gccgcctccg ccagcgcctg tcgccgtcac tctccaccca 1560
gccagtacct cgcgccagga gaacgcgccg cccgcgccgg cctcccaata gccgccgccc 1620
gctgcctgca ctctaagcgc tctccccacg tccgctccgg gagccatccc cggtcgcagg 1680
agtccgggga gggccgggag gcgccatggt ctctctctgt agcctttcct ggtagtcaat 1740
ttgttgtccc gaggattcat ctttttcttc ctcctacctt ctgttttttt tttcatgtat 1800
gcattggtct tgcctgagct cttcctcaaa gcttggagga acgggggagg gcccgggaat 1860
gtccctgtct gcacgacctg ggactctgcc catgtgcttt tgcctgtgga atggagacgc 1920
ggaccctgga tagtggttct atgactctgc gagggacagg cccacgcgtg tggggagaac 1980
atctccttcc ggggctgccc tcaagagctt ctgaaaaact aatgactctg ctgccttctc 2040
ctttgtcttt gtttcactct gtgtttccaa taaactcatt gtagcgaatc aaaaaaaaaa 2100
aaaaaaaa 2108

<210> 369

<211> 1883

<212> DNA

<213> Homo sapiens

<400> 369
gttccatcct ctgccatcta ctccactgtt cagacacctc ctaacctccg tcatgacctg 60
tggcttcaac tccataggct gtgggttccg ccctggaaac ttcagctgtg tctctgcctg 120
cgggcccccgg ccaagccgct gctgcatcac cgccgccccc taccgcggca tctcctgcta 180
ccgcggcctc accgggggct ttggcagcca cagcgtgtgc gggggcttcc gcgccggctc 240
ctgcggacgc agcttcggct accgctccgg gggcgtgtgc ggacccagcc ccccatgcat 300
caccaccgtg tcggtcaacg agagcctcct cacgcccctc aacctggaga tagaccccaa 360
cgcgcagtgc gtgaagcagg aggagaagga gcagatcaag tccctcaaca gcagattcgc 420
ggccttcatc gacaaggtgc gcttcctgga gcagcagaac aagctgctgg agacaaagct 480
gcagttctac caaaactgcg agtgctgcca gagtaacctg gagcccctgt ttgctggcta 540
catcgagact ctgcggcggg aggccgagtg cgtggaggct gacagtggga ggctggcctc 600
agagctcaac cacgtgcagg aggtgctgga gggctacaag aagaagtatg aagaagaagt 660
agcacttcga gccacagcag agaacgagtt tgtggctcta aagaaggatg tggactgtgc 720
ctacctccgc aagtcagacc tggaggccaa cgtggaggcc ctgatccagg agattgactt 780
cctgaggcgg ctgtacgagg aggagatccg cattctccaa tcccacatct cagacacctc 840
cgtggttgtc aagctggaca acagccggga cctgaacatg gactgcatcg ttgccgagat 900

```
caaggcacag tatgatgaca ttgccacccg tagccgggct gaggccgagt cctggtatcg      960
cagcaagtgt gaggagatga aggccacagt gatcaggcac ggggagaccc tgcgccgcac     1020
caaggaggag atcaacgagc tgaaccgcat gatccagagg ctgacagccg aggtggagaa     1080
tgccaagtgc cagaactcca agctggaagc tgcggtggcc cagtctgagc agcagggtga     1140
ggcggccctc agtgatgccc gctgcaagct ggccgagctg gagggcgccc tgcagaaggc     1200
caagcaagac atggcctgcc tgatcaggga gtaccaggag gtgatgaact ccaagctagg     1260
cctggatatc gagatcgcca cctacaggcg cctgctggag ggcgaggagc agaggctgtg     1320
tgaaggtgtt gaagctgtga atgtctgtgt cagcagctcc cggggtgggg ttgtgtgcgg     1380
ggacctctgc gtgtcgggct cccggccggt gacgggcagc gtctgcagtg ccccctgcaa     1440
cgggaacctg gtggtgagca ctggtttgtg caagccctgt ggccagctga acaccacctg     1500
tggagggggc tcctgcggcc aggggaggta ttaagtggcc caaaagagag ccaggggagc     1560
cccttctgcc tgccagacgt gccactgccc caccaccagc tgaaaacagc agcacatcgc     1620
tggctttttcc ccttgtgttc tgagaataca ccatcggctc attcccacca gtggctcctc    1680
cccacctttc atcccactgg aaaggggtct gtggctgggg aatagaccca ttccttcccc     1740
tgtctcagcc ttcagcccct cccggggaga agggccttgc ttccctggaa gaagcactgt     1800
gagactgttc cccctgcctc tctggcctct tgtctcccct tttccaataa acttgggacc     1860
tgcaaaaaaa aaaaaaaaaa aaa                                             1883
```

<210> 370

<211> 13464

<212> DNA

<213> Homo sapiens

<400> 370

```
gcaggttgca gtttcccttg cgttctgcct agcgggagat tcctatgctc ataaatggca       60
ttaataatgt tctctgcctt tatgttgtca ggaggcctgt cccctgccca ataaaaggca      120
ggagcagtaa caggcctccc atgtgagacc gtgtcatcct tcaatcctca gtccagggga      180
ctgtcccgca gaacagaggc atggcgagcc aatcctgcca catcagctct ggctgcgggg      240
tcaagaactt cagctcccgc tctgccactg tgcccaagcc tgggtatcac agctgtgtca      300
gtgccatggc ccatcatggg gtcagccctg gggggctagg ctccaggcgc ctcggaggct      360
ttggcagtca gagcctgtgt acggtggggt ctccccggat cgcggtgagt tgtagatggc      420
ccctacacag caggggcagg tttggctact gggcaggggg cctttgcagg cccagcccac      480
cccgcatcac atctgttacc atcaacgaga gcctcctcat gcccctcaac ctagagatcg      540
accccaatgc ccagtgtgtg aagcatgagg agaaggagca catcaggtgt ctcaacaagt      600
ttgctgcctt cattgacaag gtgggtctgt gagctgtttc ctgggactgt ggcttaggaa      660
ggggaaaggg acttggaacc cgatgatagg gtggttctaa ttctcctgtt actgactagc      720
acgtggctgt gagaacgtgg ctgaatcagg gtcttggttt tcaaatgttg ccatgtcagt      780
caaatgcaaa atgaatctga aaccaccaga cagaagaggg gtgtttgctg ctgctgctca      840
caggcaagca gcaagcatga tgtgtgtgca ccgtccatct tcccgatctc accggcacgc      900
agtaggacag tgtggcagag cgggccactc tgcctaggcc tgtctgacct cccgcttaga      960
ggagatgcta ccgtccctag catggagcag gagcaggatg cagaggggtc agaggctcct     1020
tgcctgctcc tccgtggtgg ggacatcata ggaaggctgt gaacccttag gacttctgcc     1080
cgaggcactt tataccctgc agcacatgca tagattaaag gcagtgagtt gagaaactta     1140
aagcatggtc cagggagagg ctgttgaatt gacttttctg atgctgcagt ttcagagtcc     1200
tcctggttga gtgtatagct ctggtttcct ttaacctaat gatgccctcc ctaggagctc     1260
acagtctttg gggaggaca gccaccaaag tttaaaggag gagaagctcg tcttccaggg      1320
ctgaatggtg tgtatatcta caccctggga ctgcataaga aggcacaaac cagggcgatg     1380
gtgggtggaa gttctgctct cagatcaacg ctacatcaga agtgggcata tctgcacaca     1440
gaggtgctgg gtgtccattg ttaacatacc gctggctgtc cagattactc ttcccttcca     1500
tcagcacaaa gaagctcact gtagagttag gcaggttatt cattgcacaa ggacacctag     1560
caagagcagg ggaggttgaa atccagccca ggctttgctt gcagagccat gtgccctggc     1620
agggggctgg aggcatcttt tcctaatttg ctcaaaagca cagtgagaac cctgagagtg     1680
catgctcaaa ggggtctctg catgtggcct tttgggcatc acccatgaat gctgctctct     1740
ccgttgagag gaaccagttc tgggtctggg gagaggaact tataccactg caccaggctg     1800
gcttttccct cccttggggt tgctgatggc caccctcttc cctccagcct tcttcctgca     1860
ggctccaaac tgcattgagc tccagactac aaagtgcatt gatatggtgg gtttaatcct     1920
gccagagggc tgtcccatga cctgggtctg catgttgaaa tccaaggcag atgccttttg     1980
gataatcaga gtttttgttc cttgtccctc aagcacggac aattggacag ttgagggttt     2040
gtaccaggct tgtactccca agaggttagg gttggagaag aaatcacctg ctttagggct     2100
ggtaaatgca aacttcaacg gggcttggt gggtcacatg aatgaagcca aacttgagag      2160
gcatgcccac ctagagggac agtagctgct gctcactcca gccacaggcc tacctagcct     2220
tgccagaact tctggtatgt caagagggtc agagagatcc agattcttac ttgaaaccac     2280
```

tcagttatta agtgttggca actaatcaga tctctataaa acagtgtgca aggcaagtaa 2340
gacatgcctg tgggcctgat agggcctgga gctgccagtg taccactgct aatgtatttc 2400
tcctgtcttg ttttacaatg caattcagtg agaatgaggc aaagtggtct ggttggtggc 2460
agagttagag tgagaggtct cctaattcct ggtctgggat actctgtcct gggcttcagt 2520
gccccagcac cagaggtaag gaagagtcat ggtatgtaca tttaaaaatt gctgttgtac 2580
aaatgcccac aaacttagca acttaaaaca acacaaattt actacctaaa aggctgatgt 2640
taaggtgtcg gcagggctgc atttcttttt ggaggcccca ggggagcatc tgtttccagg 2700
gacatttagg cttttggcag aattgaggtc cttgtggttg caggactgag gtccctgttc 2760
tcttgctggc tggctgccag aggttgttct cagcttccag aggctatcca cattccttgg 2820
cttgtggccc cctcttcctc catcttcaaa gccataatag cgggttgagt ccttctcatg 2880
tttcaaatcc ctctagtgca ggcacctttg gggggatcat tatttgccta gtatgcatag 2940
tatgaaatat ttctgcagga tgttaatttc attctgtgta gaataaacga acgtttataa 3000
atagtgtgct ctttctggga agggggtgg gtggagagaa gttacttggg gactacttgt 3060
gacctgcact ttcagacttt gccaataatt tggcctcctg gggaaggtca caggctgctg 3120
cagggcgggg gccatcctca atggcagaaa gagctctggg ggaccaggca aaacctgctc 3180
tcaggctttc ccctcagtct ctttcctcct tcctgaatgg caggtgcgct tcctggagca 3240
gcagaacaaa ctgctggaga ccaagctgca gttctaccag aaccgcgagt gctgcgagag 3300
caacctggag cccctgttgc agggctacat ggcgactctg cggcgggagg ctgagtgcgt 3360
ggaggccgac agtggcaggc tggccttaga gctcaactgc acgcaggagg cgctggaggg 3420
ctacaagaag aagtgagcat gacagggctg ggatggttct aggcagtgga gtcagtctta 3480
gctgagatca cacaactcct ccagaagctg agactgctga aggccagggt tgccctgaca 3540
agctatcaga cacttctgct ttaccgtctg ctttatctcg tgtcagctcc tagttcttac 3600
acccctgtgc tgggggctgg gaaaagattt agagaagagt tccctgtctg ggagggatgg 3660
gacacacatt cagggtacag actgatgcag agatgaagag ataggaacaa atgaaacagg 3720
aggagcaagg acagcagggc tggagaggag gagcccggag cagtcagagc taggacgggg 3780
ggaaagggaa aggccctgga gaggtatctg aaaatgaatt ggctgaagcc attccaagta 3840
tctgatgaac ttggattctt cctttattta ttctgcagat aaaccatcag gttatcgttc 3900
tgcagggctc cacaattcaa ataccaatgt aaatttgggt tttcctgagt tatttagctt 3960
gttcattacc tagtctaaac tcaccgggtc actgtactta ccttgtagat cagacagagt 4020
gaagggtgat tttgacctca atcaatcctg gtaattccaa ccctgcttgg gattcttgga 4080
agttcaggtt ttaccatctg gctggaccca gcagtgggtc ctcatgtccc ttggggtgaa 4140
gcaaatgtag acacaatcct cagtgtgcag gcgtcaaggt ctgtctccca ctctgggggt 4200
gtgcgggctg gtaattatag agctgagtgc ttactgtagc cccaaatgac catggaattc 4260
aaaagttatg aaaacaggcc gggcgcagtg gctcatgcct gtaatcccag cactctggga 4320
ggccgaggtg ggtggatcac ctgaggtcag gagttcaagg ccagcctggc caacatggtg 4380
aaaccgcgtc tctactaata atacaaaaaa aatttagctg ggtgtggtgg tgggcgcctg 4440
taatcccagc tactcaggag gctgaggcag gagaatcgct tgaacccggg aggtggaggt 4500
tgcagtgagc cgagatcacg ccactgcact ccagcctggg caacaagagc gaaactccgt 4560
ttcaaaaaac aaaataataa taaagttatg aaaaccatgc tctgggctct gggaactctg 4620
gtcttctttg tgggggcagg gctgctcctg ggctacaatc tcacaggcag gaggaagagg 4680
gctgggctgc tggaggaggg tgggggaatg gcttgaagaa ggggttgctt ctaggaattc 4740
aagtccttgg atattggggc aggggaata aaacctcatg ggggtggctt catagagctt 4800
tgtgttttgt acagtgtaga atttaaaggc tctgtggaat atctgtctgt ttttaaatag 4860
ttccacaaat atttattgag tgactcactg cacttcaggt gatttagacc aagacacaaa 4920
gagattaaat gtcttgcttc aggttttgca gcaacttcag ttaagctggg accagatacc 4980
agggacctgg acttcaagtc cagtgtcctt ccaatgggtt ctttgtggct gggaaataga 5040
agtccagaga ggcaaaacct cttctccaaa ggcacagaga aaggagaaaa ccagggagtt 5100
tggcacacct gggctctggc ctggccaggt cctcagtccc aggggccttt ccacctgggt 5160
tgtttctcct caggtatgaa gaagagctgg ccctcagggc cacagctgag aatgagttca 5220
tgatgctgaa gaaggtgagt aactgttctc acatggaag aattaaatcc tggacaccgg 5280
acgctctgat gctgtgggcc ctgtcaggtt cccattagcc caattgttgg tggtaaggta 5340
ctggagcccc tggcacaaga aggtgggtta gtttggttgc agaagccatc atagagccca 5400
tgcaggaatc cagacagtgc cgatagagcc ttcataaaaa gaaaacctca ccatgctcct 5460
gcttgacttt ttttaaacta ctggattggc tgaaatgttc aattttaatt ccgtcttggt 5520
tttttgcggg cattgcctcc ccctgttatc caagactagg ggttggttct cagcttgtta 5580
atgatttcaa agtgtgacca tcattcagtg caattggtga tgatggaaaa gccctgggga 5640
ataaaactgt gaatggtttc ccttttgttt catgttgctg agtgaaagtg catgtgaatg 5700
ttctttaaca agtaatgtgc ttgcttgcca gcctccttat ttgtgagtat cacagggtgg 5760
agtgaaggct ggtgacagat ctgccaatag aggcttggcc ttggtctcag gcaggagagt 5820
cagctatttg agtgctggtg gtttcctggc ctgggctctg aggacaacct acagagaggc 5880
tctgtgactg tcgcatgctt gtactctgtg ggtcagtctg gctgggaggc agcagcctga 5940
cgttctgggc cgtggtatga ggcttccttc tttggcaagg tcctcggctc tccacttcat 6000
ttccccaagt ggtctcattt atgcaatggg aggtaacgtc actggcagac aacaaaggct 6060
ctgtcccttc tagggggtcc ctgaaggtga agatcagaag tctaggttag acttgtaagt 6120
gaagggtggc tgggcagggt caggccagag gcagagaggt gggccaggag gtggcctcag 6180
ggcccgctga ggagagccga ggctcagggg cacagcgagg gggacaaagg tcactccaca 6240

```
tcctcttgtg tctatctcca ccaggatgta ggctacacct atctgcacga gacggacctg   6300
gaggccaacg tgagccgctg aaggaggagt cagtcttcct gcagtccctc tacggggagg   6360
taaatctctc cccatctctg gaagagcaag tgagagagga gacaggggag ggtttccttg   6420
gagtgtggcc tgtcctggtt cctgggttct aggaaagctg tacattcttg actctcctag   6480
gcctggatat ttttccttat gtgaccagag gggatttttc tcttgttcgg gtggaactgg   6540
gtgatgccct tcctttaaga tgatttgaaa gaaagtggtt aattcccatg tcatgaagtg   6600
gagatctgac aattaattta ttaatttagg tagtgtaaaa acccaagcag ataatccagg   6660
cctctccctc tggccacaat ctggttgatt agatttgtgc atagagctat agctgatctg   6720
gattattcag gccccaaggc ggggtgagcc ggggtgacgg catacacata aatccagctc   6780
tgttgatccc tgtattagta gaattccacg ttactcattg ctttgaggcc acagtcccaa   6840
ttccagcctt ttcctcttag aactacagtt tttaaatttt cctgacttgc tggtagtttt   6900
ctaatttcag agttgggata tttttttttt tttttaattt cctgagaaga taggatcaga   6960
aggggatttg gagttgaggg cactggttgc tgtcttctac aggaaatctg cctccttcaa   7020
tcacaaatct ctgacacctc catggtggtg aagatggaca acagctggga gctcaacatg   7080
gacttggttg tggcttagat caaggctcag tatgatgata ttgccagctg cagccgggtg   7140
gaggctgaga cctgatacca aaccaaggta gtctggaggg cagggcagcc ccactgagac   7200
aggaggatag ctgtgagcct cgagtttccc tgggaaacgc actcctgcac tctagagggg   7260
ctgccgctga ggtcctgggt agagcagggc cagaaggcgg gagagctgtg gaacagtagg   7320
ctggctgatt gggtcgggga cagcagcctc tagccattgt ttctgcgtgt ttgctctcaa   7380
cctgcagtgc gaggaggtga aggccacagt gacccaacag ggtgagaacc tccgcagaac   7440
caaggatgag ctcaacgagc tgaactgcat gatccagagg ctgacggcag aggtggagaa   7500
cgacaagcag caggttgggg ggcgccgaga ccccgctcca tacacctgcc acttcttgtc   7560
tggttcctgc ttccctccta tcttcaggga ctgctaccat ttggagaggc gtgtatttga   7620
actgtgccag agtacactac tataatcagc atgcaagggc tcctgttcca acacccacta   7680
ttatcaaact tccaaattct tgtggatctg ataggtgtaa actggcacct aattgttatt   7740
ttaattttca tttctttggt taagtagtga tatcaaatct ttcctcatac acttactagt   7800
tatttgtatt ccccttctaa gacctgtcta ctcatatatc tgcccatttt tctattgtgg   7860
ttccttcctt ttttcttagt gatttgcagt attcctttgt ccattctatc tggattgtaa   7920
gcctttgtca gtacacatat tgcaattaac tttctctcag tctatcactt gtctgtttgt   7980
cctttattga gcccaaatct ttaagctgga tgtgcccaaa tccatcttaa gttcttcctt   8040
taatggttta gacttagaag gtattgaggt aatagaggta gcctcttaaa tttctccccc   8100
cgttagtttt atagttttac ctcttgcatt taggtctgtg atccatctga actttatttt   8160
tttttttgt atacacagtg agtagggttc taaattaatt tttctacata tcataagaaa   8220
atagattatg taatccatcc tttcgccagt gatttgtaat actacctcca gcacaacaca   8280
gtgctcccac acatgcctgg cttaatctct gtcttctgcc ttctgagttg ctgttcttca   8340
tcactgcaca ggatgtatca ttagtaattt gttgtggtta agacctgagt aatgaaatat   8400
gcgagtggtt tactttttag atgtgaaatt agagaagatt cgaagttgtc ccgtgataac   8460
agtgcttttg ttcatttgct tctgatatgc tttcttaagc ggggtgaagg gaggggctgt   8520
agttagtgtg tccggagttg ttgcacatgg gagcttcttt cagcaagatg atctcagcag   8580
gccaaagagt gagaccctgc acagtggcct ttgaatacct gagcttggtg ctaattaact   8640
gcatagctta gggcaagtca cagcttttg gagtcttaat ttcttcattt gcataatggg   8700
aacaataatg cctgtagtga caagcttgca gaataacttt gaaggtcaaa taagataaca   8760
tctatgaaag taccttgtaa acgaacaacg gtatatacaa gtataaggta atatcactaa   8820
taatgatcat tcctatttca gtgctaacat ttcctccctc taatggctca cctgccaaac   8880
tcaagtcagt tttatgataa agtgtacaaa aatgaaggaa aaaaatccaa atgggccatg   8940
tccctgccta gttaatttc catttatatg tatttctttt aaatagtatt aagtttaaaa   9000
attaaccgtt ttgaatggag ataaaatagg atgaccaagg ccactggaga cccaccagag   9060
tccccacaac ctgcaccctt tttaaactcc ttcccttctg tccatcccca cagcgctgca   9120
agctggaggc tgcggtggcc cagtctgagc agcagggtga ggcagccctc agcaatgccc   9180
actgcaagct ggctgagctg gaggacgccc tgcagaaggc caagcaggac atggcctgca   9240
tgctcaagca gtaccaggag gtgatgaact ccaagctggg cttggacgtg gagatcgcca   9300
cttatcgcaa actgctggag ggtgaggaga tccggtgagg acaggggctc cagggtccct   9360
tcaggttcct actcagagct ggactgaaat attccaggca gagaatgttc aaactgggaa   9420
agactctaga aatcatggac tccaaccatg tccccatttt gcagataggg aaactgacat   9480
cagagagggg aaggggcttg cccaagagca cagccagtta atggctcacc taaaatgtac   9540
cagtctctgg gctggccacc aatgctcttc tccctcctgc acggcccctc aggctggaat   9600
ctggatctca ttccaaggcc actactttac cctcttgact gaaaggatgt ttctgcaatc   9660
agcttcagtg ggttcttaaa gctggaatca ttcagaaaat tagttaaagg ccatttgatt   9720
tgggatgtgc aaagaatcct atcaacctag atagcctggt gttcaggacc agggaataga   9780
aatgagcggc acaagttctg tccttgcaaa ggaagacaag acagacatac atcagtcctc   9840
cagagcctgg ggttagtggc ccagggaagc tgtcacagaa ggaggctgcc ttctcaacag   9900
ctgaggagaa gctatgctga caatttgagc caagaattga gggcaggtga agccagtgtg   9960
gggcagagtg atgcgtattt acacctggga tggttgcatg gtcctctgga cttgggacgg   10020
ccacatcgtt agagtttcta gctgggtccc ttcagggcgg ggaggctaga ctgggtagta   10080
gagtatacag gtgaagagca caggccatga gcccactctg ggttcaaatc ccagctctgc   10140
cacttattag catgtgacct ctttgtacct cagcatcttc agctgtaaaa cagagataat   10200
```

```
cacaaaatcc accttgtaag gctgttgcag gattaagtga gttttaatgg aacatagtaa  10260
gccccagtag gaggaggtga ttgtcatgag atgatggcca ggctgggctg cccccatgaa  10320
gagggaactc aaaggcctgg ttggctttgg aaaccttcca tgaagcgtca ctgatggcag  10380
cccggccacg ctgctggact tgctccatgc ttggctcaag cacacacact ttctgccag  10440
cttccttggg gctttgggtg aagccactgg agttctgtct gcattcggat acttaagccc  10500
ccaacccact cactagcatt cagtgagcag ctggcacttg cactgatttt cattatgata  10560
ctaataatta atggattaat aactcaatca gaattattag tacacatccc ctatccatgt  10620
gattttaaag tcaatttagc atattgtgct atgtgtactg tggccattga ccttttaaga  10680
gactgcctcc tcttggcatt taaaatattt aatttaatta cataagtaat gctcgctgct  10740
gccaatacag aaaataagaa aacaaaagca ttcagagcca ctcaactcag agatagctgt  10800
cgccagcatt gtggtttgac aggcaccacc acgggcactg tgggttcctt gaaagggcac  10860
agtcgtaagg ctgtggcaat gccagactga gtggcaggaa accgagccag atgtgtacat  10920
cctcactcaa aagatgagtg aagaagtgtt ttgagctggg cagaatgagg cagggtagag  10980
ctttcctggg gaaaaaggtt tgaagtggga agctacaaaa ggggagacag aacagaaagt  11040
gcagagaggt gtgagggagg agggagtgtt tataaaaagc ccactgtagt ggtatatatg  11100
tctcatcata aataatatcc acatattttt tgagacagag tctcactctg ttgcccaggc  11160
tggagtgcag tggtgcaatc ttggctcatt gcaaactcct cttcccggat tcgagtgatt  11220
ctcctgcctc agcctcccaa gtagctggga ctacaagtgc ctgccaccat gcctggctaa  11280
tttttgtatt tttagtataa aaagggtttc gccacgttgg ccagcctagt ctcgaattcc  11340
taggctcaag tgatccacct ccctcagtct cccaaggtgc tgggattaca aaaatgaatt  11400
gtttataatg ttggtagcaa acatattgga tttccttact atgtgccagt ctccatttta  11460
ggcactttcc atgtagaact caatcgaact ttacaataat cctttcaggc aggtattatc  11520
acccccattt tatgctggag aaattgaggc acagaaaggt taagcgacct gcccaagatc  11580
acccagctag taagccgcac agccagattt aagtccaggc agcctggctt gagagtccag  11640
gctcttaacc acaccatact cttcagcctc cctctgcatt gaatgaaatc agatagatgt  11700
ctggggtacc atttaaaatc atggtataaa aatgaatctg tttgagtaat aggcaagctt  11760
tttcctgctc agagacgtct agctaggtag cagcagcttc catcaccact ttggctttgc  11820
tggcatttcc agccaaggga ccaggaggat gccccctca tctctacttc ccagttcacc  11880
aaatttccat ggctgaaagc caaacatggc ataaagtggc attttggatc ctgtaagcta  11940
acatactcct gtcctttgg gatcccagca agcaggacca acactgatgc tgtctgcagc  12000
aggtgggtag cttcccctg cctcttaggg cccttcttct ccttcagcat ggaggagaat  12060
aggtttggtg caaaagcaga ggaactcagc actaaccgta cactctctcc tctccctgca  12120
ggctgtgtga aggcgtgggc tcagtcaata tctgtaagtt gggtgtgtgt gggcatggct  12180
ttctggggtt tcagcaatct tagtattaga ttgcatctaa ctgatgacca agcctctgtc  12240
tatggagtgg tagacaaggg actcctgagc tgatgtctct gggatccaga caaactgatg  12300
acctcaaata tttgaccata catggatcca tctggaacac ccatgaagtc taagttaatg  12360
aatggtgcaa attctgaacc cctatttggt tctgcaaaaa tagcttaagc cttttagtta  12420
cacacagaat actactcaag tggtagcatg ttctagtaag aaaaaaagct aggtttatat  12480
tctgatcatg tgggtctctg caggcaactt gtaaatctct atctaataat ccaatgtaca  12540
cacagcccct gggccaggtg tctgattccc atccctccat tcccctcacc tctctagcaa  12600
ccatatttct agatgactcc tctagtgtga tctatcagtt tttgctcaaa gattgtatga  12660
acatgtctca acatcagcct ggtccaagta gagtttggtg tgcaatttgc taaccacact  12720
cctgaagaag cagatggcag tgggttcagg tttgcctgta gtttaaaaag actcatgttg  12780
ctggtccctc agagaaggaa gataattaat cattggttct aaaggattct gggtcataac  12840
atgtctgaga tggaagagaa ttctgacctg tgtctctgtc cacatgtgtg agccattccc  12900
agggtggtgt ggtctgcggg gacctggact ccactgcctc ctgtggctca gggggtgtgg  12960
ccatcagcag tggtgcactg tgttgcccct ctgcagtggg ggcctgctcc agtgcccaat  13020
ctgtgcggtt tgcatagagg gggtgggact ctggacggac ctgttctgat gggggcagtg  13080
tgggttgggg agtggaaatc ttttctcttg catgaccaaa tgtcagttta gtaatttccc  13140
cttacagaga acccagaact aactctcctg tttccttctt attggagtta acaatttact  13200
ttctctgtgt aaaccctggg tttgacatta ttttcacttc cattgcttct ccccttccct  13260
gtggttcagg aatgaaatag aaaggctttt ttttgaatga atttctcggt cattgtcttc  13320
cccaattaac ttcaggccag cagatgttaa attaagtcct gactgtgtgc cagatcctct  13380
tggaacggag ccagctagag ggtgtggagt atgtgggaac atctctgcgc tttgccctcg  13440
gatgacaggg gctccagatg ctgc                                          13464
```

<210> 371

<211> 2508

<212> DNA

<213> Homo sapiens

<400> 371

agctctcccc accaataaaa ggaccaggga ggatcagaga gagcagaagg atcctgagcc 60
tcgcactctg ccgcccgcac caccttccgc tgcctctcag actctgctca gcctcacacg 120
atgtcgtgcc gctcctacag gatcagctca ggatgcgggg tcaccaggaa cttcagctcc 180
tgctcagctg tggcccccaa aactggcaac cgctgctgca tcagcgccgc cccctaccga 240
ggggtgtcct gctaccgagg gctgacgggc ttcggcagcc gcagcctctg caacctgggc 300
tcctgcgggc cccggatagc tgtaggtggc ttccgagccg gctcctgcgg acgcagcttc 360
ggctaccgct ccgggggcgt gtgcggaccc agccccccat gcatcactac cgtgtcggtc 420
aacgagagcc tcctcacgcc cctcaacctg gagatcgacc ccaacgcaca gtgcgtgaag 480
caggaggaga aggagcagat caagtccctc aacagcaggt tcgcggcctt catcgacaag 540
gtgcgcttcc tggagcagca gaacaagctg ctggagacca agtggcagtt ctaccagaac 600
cagcgctgct gcgagagcaa cctggagcca ctgttcagtg gctacatcga gactctgcgg 660
cgggaggccg agtgcgtgga ggccgacagc gggaggctgg cctcagagct caaccatgtg 720
caggaggtgc tggagggcta caagaagaag tatgaagagg aggtggccct gagagccaca 780
gcagagaatg agtttgtcgt tctaaagaag gacgtggact gtgcctacct gcggaaatca 840
gacctggagg ccaatgtgga ggccctggtg gaggagtcta gcttcctgag gcgcctctat 900
gaagaggaga tccgcgttct ccaagcccac atctcagaca cctcggtcat agtcaagatg 960
gacaacagcc gagacctgaa catggactgc atcatcgctg agatcaaggc tcagtatgac 1020
gatgttgcca gccgcagccg ggccgaggct gagtcctggt accgtagcaa gtgtgaggag 1080
atgaaggcca cggtgatcag gcatggggag accctgcgcc gcaccaagga ggagatcaac 1140
gagctgaacc gcatgatcca gaggctgacg gccgagattg agaatgccaa gtgccagcgt 1200
gccaagctgg aggctgctgt ggctgaggca gagcagcagg gtgaggcggc cctcagcgat 1260
gcccgctgca agctggctga gctggagggc gccctgcaga aggccaagca ggacatggcc 1320
tgcctgctca aggagtacca ggaggtgatg aactccaagc tgggcctgga catcgagatc 1380
gccacctaca ggcgcctgct ggagggcgag gaacacaggc tgtgtgaagg tgtgggctct 1440
gtgaatgtct gtgtcagcag ctcccgtggt ggagtctcct gtgggggcct ctcctacagc 1500
accaccccag ggcgccagat cacttctggc ccctcagcca taggcggcag catcacggtg 1560
gtggcccctg actcctgtgc cccctgccag cctcgttcct ccagcttcag ctgcgggagt 1620
agccggtcgg tccgctttgc ctagtagagt catggagcca gggcttcctg ccaagcacct 1680
gcctgcctgc atcactgcac tgaatggcat gtgaatggaa aatgtgtgct tgcttccaga 1740
atcttctgga tgttcctaca gagggaaaga cctacagagg gaaagaccct cgggccgctc 1800
ccctgcgcct tttcatgcta gggagatgca tcctagttgt cctcctggca gctgtttca 1860
gaggcattcc cagcccttca cttaactcct acttagctcc aaaatacctg tatccaattt 1920
gtattattcc cccagctctc agggacaaga ccagtccccc agcgtggtgg tcagcacgga 1980
agctccacct tctgggtgga ggcgccatcc taaccatcca gccaggccac ccacaacccg 2040
agaatcaggg agaaagtccc tccccagcag ccccctcctc ctggctggga agaatggtcc 2100
cccagcaagc acttgcctgt tcattcccgt tcatgttttg cttctctctc agactgcctt 2160
cctgcttctg ggctaacctg ttccagccag gctcctcatg tgacctcgca gttgagaagc 2220
ccattatcgt ggggcatcct tttgcctaca gcccctggtt agggcacttt ggacaggtct 2280
tgctattcag tgaacctttg tacatttcaa agaagactcc atggctgctc cagatgcccc 2340
cttgctgggt gcaggtgggg actgtccaat gcagagctgg cgggacagag agttaagcca 2400
cttcctgggt ctccttctta tgactgtcta tgggtgcatt gccttctggg ttgtctcgat 2460
ctgtgtttca ataaatgccg ctgcaatgca aaaaaaaaaa aaaaaaa 2508

<210> 372

<211> 2404

<212> DNA

<213> Homo sapiens

<400> 372

gctcacccat tctacacttg ctcttttgct cccaaccagg aagccatcat gtcttgccgc 60
tcctaccgag tcagctctgg tcaccgggtg ggcaacttca gctcttgttc agcaatgaca 120
ccacagaacc tgaatcgctt ccgggccaac tctgtctcct gttggagtgg gcctggattc 180
cggggccttg gcagctttgg tagtcggagt gtcatcacct ttggatcgta ctcaccccgg 240
atagcagctg taggctctcg gcccatccac tgtggagtcc gctttggtgc tggctgtggg 300
atgggtttg gtgatgggag aggtgttggt ctggggccta gggctgacag ctgtgttggt 360
ctgggctttg gagctggcag tggcattggc tatggctttg gtggccctgg ctttggttac 420
agagttggag gggttggagt cccagcagcc ccatctatca cagctgtgac tgtgaacaag 480
agcctactga cccccctcaa cctggagatt gaccccaatg cccagagggt gaagaaggat 540
gagaaggagc aaatcaagac cctcaacaac aagtttgcct ccttcattga caaggttcgg 600
ttcctagagc agcagaataa gctcctagag accaagtgga gcttcctcca agagcagaaa 660
tgtatcagga gcaatctgga gccactcttc gagagctaca tcaccaacct gcggaggcag 720

ttggaggtgc tggtcagtga tcaggcccgg ctccaggctg agaggaacca cctgcaggat 780
gtcctagagg gcttcaagaa gaagtatgaa gaggaagtgg tatgtcgggc caatgctgag 840
aatgagtttg tggctctgaa gaaggatgtg gatgcagctt tcatgaacaa gtctgatctc 900
gaggccaacg tggataccct aactcaggaa attgactttc taaaaacgct ttacatggag 960
gaaatccagt tgctgcagtc gcacatctca gagacgtcgg tcattgtgaa gatggacaac 1020
agccgtgacc tgaaccttga tgggatcatt gctgaggtca aggcccagta tgaggaggtg 1080
gccaggcgca gccgggctga tgctgaggcc tggtaccaga ccaagtatga agagatgcag 1140
gtgacagctg gccaacactg tgacaacctg cgcaacatac ggaacgagat caacgaactg 1200
acccgcctga tccagaggct taaggcagag attgagcacg ccaaggctca gcgtgccaag 1260
ttggaggctg cagtggccga ggccgagcag cagggcgagg cgaccctcag tgatgccaaa 1320
tgcaagctgg cagatctgga gtgtgccctg cagcaggcca agcaggacat ggcgcggcag 1380
ctgtgcgagt accaggagct gatgaatgcc aagctgggcc tggacatcga gatcgccacc 1440
tacaggcgcc tgctggaggg cgaggagagc cggctctgtg aaggtgttgg accagtaaac 1500
atatccgtca gcagctcccg gggcggcctg gtgtgcgggc ctgagccttt ggttgccggc 1560
tccaccctct cccgcggcgg ggtcaccttc tcaggtagca gcagcgtctg tgccaccagt 1620
ggggtcctgg cttcctgtgg ccccagcctg ggtggagccc gggtcgcccc ggccactggg 1680
gacctgctga gcactggcac aaggagtggc tccatgctca tcagcgaggc ctgtgtcccc 1740
agcgtcccct gcccctgcc cacccagggg ggcttcagca gctgcagcgg cggccgcagc 1800
tccagcgtcc gctttgtgtc caccaccacc tcctgccgga ccaagtactg agagcccagc 1860
cccagacagc tgctgcccag agaagaacca gctccacggc tcctgcttct gcccccaggg 1920
ttcgtgggct ctgggcttga cggtctccag ctccccttc tgccaggaag ccacccctta 1980
gcactccccc gattattctg cctggcccca tgtcctcttg gaggattttt ctgccatgta 2040
gatgctccat tagcagttcc agctaagctg gcttctccct gcctcccaat ttctgtcctt 2100
tggatcactc ctctccctaa accccgagat ggcttttctc ccagtggctt ctctccggct 2160
gtttctcttc ctgggttgtt ggtgtaactc cagcatgcaa gtccacgagg tggggtgcag 2220
gcggacaggc gggctttgca atgatccccg tactcctgcc cctgctcgtg gtgtgaggcc 2280
cagccctgtc cactgcaggg ttgtgtctgc aagctccaag gaggcctggg ggttgggtgg 2340
ggtgggctgc ctgtgtgcac tactctttgt gttcctgggt tttcaataaa cttgccaagc 2400
tcac 2404

<210> 373

<211> 2681

<212> DNA

<213> Homo sapiens

<400> 373
acatctcagt cctccatccc cctaccagat aaaagggggg aagctgagcc tgacttgatc 60
atcctgcccc ggttctttcc tccagggccg catttctctg cctctctctc ctgccatgtc 120
gtaccactct ttccagccag gctccaggtg tggcagtcag agtttcagct catactcggc 180
tgtcatgccc cggatggtca cccactatgc agtgagcaag gggccatgcc ggcccggggg 240
tggtaggggc ctccgagctc tgggctgcct tggctcacgg agcctgtgca acgtgggctt 300
tgggaggccc cgggtagcct ccaggtgtgg aggtaccctg cctggcttcg ggtaccgact 360
gggagccacc tgtgggcctt ctgcctgcat cacccctgtc accatcaatg agagcctgct 420
ggtcccactg gcactggaga tagacccgac tgtgcagagg gtaaagaggg atgagaagga 480
gcagatcaag tgcctcaaca accgtttcgc atctttcatc aacaaggtcc gtttcctgga 540
gcagaagaac aagctgctgg agaccaagtg gaacttcatg cagcagcaga ggtgctgcca 600
gaccaacatc gagcccatct tcgagggcta tatcagcgcc cttcggcggc agctggactg 660
tgtgtccggg gaccgcgtga ggctagagtc agagctctgc agcctccagg ctgcactgga 720
gggctacaag aaaaaatacg aagaggagct ctccctgcgt ccctgtgttg agaatgagtt 780
tgttgccttg aagaaggacg tggacacagc cttcctgatg aaggctgacc tggagaccaa 840
cgcagaggca ctcgtgcagg agatcgactt cctgaaaagc ctgtatgagg aggagatctg 900
cctgctccag tctcagatct ctgagacctc ggtcattgtg aagatggaca acagccggga 960
gctggacgtg gacggcatca tcgctgagat caaggcgcag tatgacgaca tcgccagccg 1020
cagcaaagcc gaagcagagg cctggtacca gtgccggtat gaggagctga gagtcacagc 1080
tgggaaccac tgtgacaacc tccgcaaccg taagaacgag atcctggaaa tgaataaact 1140
gatccagcgg ctgcagcaag aaaccgagaa tgtcaaagcc cagcgctgca aacttgaggg 1200
tgccatagct gaggcagagc agcagggcga ggcggctctc aatgatgcca agtgcaagct 1260
ggcagggctg gaggaggctc tgcagaaggc caagcaggac atggcctgcc tgctcaagga 1320
atatcaggag gtgatgaact ccaagctggg cctggacatc gagatcgcca cctacaggcg 1380
cctgctggag ggtgaagagc acaggctgtg cgaaggcatc gggcccgtga atatctcagt 1440
gagcagctcc aaaggcgcct tcctgtacga gccatgtggg gtcagcacgc ctgtcctcag 1500
cactggcgtc ctcaggagca atgggggctg cagcatcgtg ggcactggtg aactctatgt 1560

cccctgcgag ccccaggggc tactgagctg tgggagcggg cggaaatcca gcatgacgct 1620
aggagctggg ggcagctccc ccagccacaa gcattagcat gatccgagac atccaggaga 1680
cagagcccct gccccatggc cctggatgtc acactcccac cagggctgaa gacaaggatg 1740
ttccaaaacc ccacctccct ttattgactc cacattcccc tccagtgttc cctccttgag 1800
agctgagctg cccctaggac ccctcctttg cctcatgagt caccttccac tcacctgtgt 1860
gcagaccctc agctaggccc agatggtggg gacagacggc agagaaagag catgcattgc 1920
tcttaaccgc agaatggcaa taacccccaa aaggtcaaat gggcacacca cagttctgct 1980
ctgtggatca tcccatggaa gtttcctcac aaagtcagct cctctcccac cacaggctca 2040
cactgccctt actttaggct ctcatgcccc taagaaaagt gaatttcctt taacaccgcc 2100
tgaaacatgc acactgcaaa tatcaaaagg agacaccccc aaatcccact ctaaattcca 2160
aagccaagtg ccacgatttc aggatcagtc aggacctgcc aatgtcctct cagcacagag 2220
aatcaagagg tggcctttgg gcagtagatt tactttctca ggtatgttga acccaaactg 2280
tgagaataaa tctttctcct gagctcacat tggcaggtga tgtcaggcca atgcctagaa 2340
caaagggcaa agggatatgg aatgtttact ggacagttgg aaaaccaaaa acagatacca 2400
ttttctcaat ggaatcagtt cccccctccac ccatcccctt ctgtacatag ccgcccttc 2460
ctgttcccca accctttagc tccttttgct ttgagtctga tatgacagtg atgtggggct 2520
gccccagctg acttggaggc cctggctcag gagctcatag atggaggagg tgtgggctga 2580
ctacctccca gaagcccctc atgccaagcc tgtcccactg gggctccagg gcctctgccc 2640
tgaacatctg ggcttctgtt caaaataaac ctactctgtt c 2681

<210> 374

<211> 11529

<212> DNA

<213> Homo sapiens

<400> 374

cagagtatat ggacaactga gtccagattc taggtctaga gagatcatgt cagaaagaaa 60
atggacatat taaactttct gccaaggacc tcctgcccat ttaatggaat tcaaaacttc 120
acccagtaat atgatggtgg cgtgttcttt tcttgaatta tcacacatgc acctggttgg 180
tgccagcccc tctgactctt cctgcttctc ccctctccaa acaaaaccat ctgtccccca 240
accccattgc accaacatct gcttggctct ctatgttcca ggttccttcc aggctgagga 300
agaggaggaa gagcctgcca cctggactct gagatccagg ttgtgcagag cctccctgca 360
ctgttaggga tcccatatcc cagagccctc tgggaggctc ctgcatggaa gctctctcag 420
tctcttacct ctgactgggg caggtgtggt ttctggaaca ggagaacaaa gtcctggaga 480
ctaaatggag cttcctgcaa ggccaaaaaa ccaccagggc caattttgag cccatgtttg 540
atgtccacat cgacaacatg aagcagcagc tggactgcct gggtggaggg tatgtgaagt 600
tggatgtgga gttaaagaac atccaggacg tggtagggga cttcaagaac aagtgagcag 660
acgctcagcc agtgcctggg ccagggtcct gccctggggc tccaggggag gtcagagggg 720
gcaggagaaa cagcctctgc tctccagaaa tgtccagtct agtggggaaa cggaacacac 780
ttagggcaca gaggtggaat cagaagtaaa cataaaagtg tcaagagaac tgggacccac 840
cctgtggtca agaatgctga ggctttagtg aagggctggg attgtagatg atgataccta 900
acataggtac aaaaagcatg atacacacag cattcttcaa actctgcaga taaggaaagt 960
aaggctctga gatagcatac caaagatgac acaaagtggc aactcaggag ttaagaatga 1020
atcccataga atctcactct cacatgaatg caactaaatc tcactgctaa tttgcacagt 1080
gttcacatat tcactctaat ccaggcaaag aggctggaga acaaaaaaga aaggacacat 1140
ttgctgcctt ctcccaggac cacacacaca cacacacaca cacacacaca cacacacaca 1200
ccactaaaca taatgagatg ctcaaatgct catgcctgta gatacaaatt tcttgatttt 1260
attgttttta cttattttta tttcttattt ttctggttct aattattttt aattctcttg 1320
gtctaaccaa aattagtaac tggatcatga gttttggtta acactgtttg tatactctag 1380
cacaagttca ggttttagag ttgtgctgtc caacatagta gccaggagcc agtggccaca 1440
gatgactatt taagtttaaa ttaattaaaa ttaaataaaa tttaaaattc agttcctcag 1500
tcacacttgc cacatttcag gtgttcaatg gccacacatg gctaatggat cctgtaatga 1560
acacagatac agagcaattc tgttatcgtg gaaagttctc ctggatagca ctatagactt 1620
cagagtcagt ggatttagga tcttggctca accacttaat ggctgaatga ccttaagaaa 1680
actatttaat ctctcaaagc cccaggttct ttatctgtta aaagagacat caatagtgcc 1740
cacctcacaa gagcattgag attattaaat caaaacttgc atgtaaggtg cttggtaggg 1800
ttcacagtgc ctggcacata gtaagttgtt aataagtgtt agctatgttt attgtattta 1860
tgattaggta cgtgtagtga tacagacgca gacaagagaa gcaagtgcag atgggtaatt 1920
tataacatgt atgcaagcgt ggactaacag tgtttcaata caagaagact tcctggagga 1980
gaaggctaaa gcttagcatt gaaaaggaag aaggcacagg ggagggatat ggtagggcat 2040
ctgttgaaga gtgtgttcct agtcaatgac caggcagtgt tcctgtggca tcaaagtggc 2100
ctgtgcaggg ccaggcgatg cttgcccttc agcagaggga ctgtgagctg ggccacatgg 2160

aagacgagga cagttgtatg gagttaggtg aagaagagac attcagagtt cattaaaagg 2220
caactcaaag aaacgatgag gaaaaccaaa ttgctcttcc cagagaacac tgagagaaag 2280
cacaatactg aaatttcaat ggcaagcaag agattttctc agatcctata gagcagcaga 2340
ggaaatataa aagccaggaa gatcctgctt catgtaagtc atgtgtgatg ggactctgtg 2400
gttttaaaag aatgacaact ttccattaca tagtagagaa agatcctctg agaggtcata 2460
tggctgatgg gcagagcagc tgagatcaaa accacagctg tctagcacca cgagggtggg 2520
gtgtgctgat gcacccaagt aggttgggtg gagaatgggt cacctggaac agaagcaaaa 2580
cttccttccc agagatcagt gagcaggagg cacctccaca tgcccctggg ggacagacag 2640
ggtggggttg ctgcctacag gtaaattgaa gaatgagaaa cctttggggg agagggttcc 2700
tatggtcctg aagtctctgc tccttaggcc acagttcagt gtgcctgaac agagtgaggg 2760
aaatgcgagt gttgtacagt gaggggctgc cggcctctga cctttctccc tgttcctctc 2820
ctccctgccc tgggttgggg caagaccata aaaaaaaaaa aaaaaaaaaa agggaaatgt 2880
atatgcccct ttatgggacc caaatactcc aaggagctga gctgtgagaa aagaaacctc 2940
aagtagactg agaggggatc tgttttgttg ctttgggttt gggttgtgtt tggtggttta 3000
tttgatttgg gatttggggc catgtgtgtc ataggagagc caggctacag tggagcattt 3060
gctaaggctg ctctgcagac ccaccattgc ctgtccacct ccaccctact gctctgacct 3120
cactaccagc gcccaatcca accccaaacc acttcttgat ccttcccaga ggtatgaaga 3180
agaactcaac aggtgcactg aggcagagaa tgagtttgtg gtgctcaaga tgagtggaca 3240
gagttcctcc tgcccagctc ccagtggcct catacaaggc cttatttaat ttgcacagca 3300
agcctagaag taggtcttac tttaacagtg ggaaaactgc aattcagaga ggtaaagaca 3360
cttgcccagg gctacctagg taatcaggag tgagtcagga cccaggactc caaagtctga 3420
gctgcaaatc ccaacaccag cccccgccgt gtggagtcca gtgtggagac aggggaagca 3480
aaacgggcct tgtcctgggc tgggtctcta ggcacccagt caggctcagg aggacagggc 3540
agcccagaag cagctccacg gagtcctggg ccttcctgct tggtggccag ggccagaaca 3600
gggttgtcct cacaatgtga ggtgatgcag gagacctttc aatgtggcac cagagctgat 3660
gggcaggtgg gcacccctga aggccctggt cttcacagag tttggaagtt ctactccctc 3720
cctcctctcc aagttctaag ttctactcca tggcctggag tagacagatg ctctgtcctt 3780
gggagctgaa gctctcccac ctaattcaca gacagagggg gctggttatc cctggaaggg 3840
taaggcccag agtgatagca tgttctgtgg gacatggatg ctgtctacat gaacaaagtg 3900
gggctggagg ccaaagtgga tgccctgatg gaggagacca acttcctgag cactttctat 3960
aaggcggtga gagtgccccgg tgccccctcc aacagggggg ctgggggctg ggtcttagag 4020
cctcagctgg gcacagagcc tgtgggctcc ttccctggtc ttctctctgc tccttatccc 4080
acctgcgtat tgcagggcag atgccacttc ccctaccaca ggagaaaatg aggttccata 4140
gagcggctgg caggtcagga gagatgaaaa attcagcctt aaccatctgg atcaagcatg 4200
ttctgttggg caggcagaaa aatggaccac ctgtcattaa aatatggatt tccatagaac 4260
tttctgagca acttggctca tcctgccctg ccaaaattta cctgctgctc ttcctagagc 4320
aactgcaggg tgatgcaatt acacaaatca agtggcaatt agtgagccct caggtgcttc 4380
tgggagcaaa cagcacaatt gcctggttct ccgtttgtgt cctccacaga gctgcaggca 4440
cagtgctgag caaatacaac ggggcctgta agttcttgaa gcttacagta tagtgcagga 4500
gaaaggcaca acagagataa tcaaatcact gtactatttc aggccctaca gaggcaatga 4560
agagaaatga cagtagaaag caatcatagg cagtgggtgg gtgacactgg acaagaaact 4620
tgaaggaatt gagcatgtca gccaggaagg catggtgggt gggggcagtg ctccagcaag 4680
aaggaatggg aagtataaat gcccccgggt gtgatccagg aatgccagga ggaggacagt 4740
atggctgggg cagaggggag agtggtgaag atagggccag cagtgtcaag gtggggtgta 4800
gggtgagtgg agtgcctcca ggtaggcctg gtaaggacag gtgatatgcc catagggtca 4860
cacacacgtg atgcccaaca gaagctggtc acattgttcc aggacctggg gcttcagctc 4920
tgattgtttg ggtgggctcc aattccacct tacaacccac agggggataa ccagcccccct 4980
ctctacttgt tctcttcaat gctgacaaca gggtgatatt gcctccgaat tctgaactag 5040
cccacagcca gcaagcctgt gggacaagac ctttgaacta gatctgctgc caggaaaaac 5100
tatttgggaa aaacaagccg cagcacaata agaaaataat aatttgttga aacggaaat 5160
gtatgaatca gtcatcaata aagagagaaa aggaatattt ttacttagac ttctgctttg 5220
gcctactgat tctcaaattt taaagagcat cagaatcttc tagaagtgtt gctaaacacc 5280
aattctgatt tagtaagtct gggctgggcc cacagtgttg catttcaatc aagtttccag 5340
tgatgctgat gctgatgctg gtttgaggtc acattttgag aaccactgag ttagatgatc 5400
ctgaaaatgg accattaaca aattaattgt gggttcactg tttccagacc atcttatctg 5460
gatgtcattt aatctccaca aattcatttg aggccagtat tatgatcagc cttgctgaga 5520
tctgggatgc tatttaaatg tgattggcag atacaagact caaatttgtc agttgggtgc 5580
cctttctgct atttccagct gatttccatt tcccatccta tccccgagga gcagggagga 5640
gccaaattgg gaaacgcaac tggaaagtag ccaagttttc tggcttcttc tgactcagtt 5700
ttctttcaca agatattttc cttatttggt gtgaagggct gtattttctc agaacatgag 5760
acatctgtga tactgacagg tcacaacatg gtgtgctcct gtgacaacag agtggcttct 5820
ggagaatgct ggctggagtt acatgtgctg ctgctccctc cacgcagatg agctgatgac 5880
aaagaacagc tataaactca atgacctccc taatgttctt cccagccctt tcagcctaca 5940
caggccactc catccctacc gtgcaagacc ctcaatgcca gggtctgttc caggggggcag 6000
atcacagcct cccaaaacat gcccccacac acacatgcat acacatgctc acacatgcac 6060
atgcatgcac acacatacat gcactcgcat gcacatacga acgcacatgc acacttgcac 6120

acacacacac ggacacacat gcacatgcac actctgcctc agctctctca tttggaccca 6180
accacaaaag ctggctcagc ttcaggccca gatctctgaa acctccatgg tccttctgtt 6240
gatggaaaac aactgcaacc tggacctgga tagcatcatc attgaagtca aggcccagca 6300
tgaagacatc gccaacaaca gcagggctga ggctgagttc tggtactaaa ccaaagtgag 6360
acatcaaggt caagttcatt aaacaagcac aaaaaaccga gatttcaaag gcaaacaaga 6420
taaaagccat gtcctctcag atgtcccagt ctgatagggt caagggatct gagttcagct 6480
tggggctcac cctatgttac ctgggggaag catgtcagga cctggggaac agccagataa 6540
aggcaggagg tggtcagtcc ctctgccaag aggaagatgg acacacctcc aatgccctgt 6600
atcagagaga aaagggacag ctggacagtt ccagccatca gtgccaacag gaggaacaga 6660
gagggagggt ggggacaaga ggaaagttta acccсttgtg gaggaggctg gaggatccag 6720
agaagaaagg cggaactggc tggggcccag gggctggagg gcaaagtctt aatctaagtc 6780
acagactggg atttccaga tgagagcaaa cattctgccc tgggagtcgt tgggatgata 6840
atttgcaatg ccgaggaaaa cacgaacaaa ggaaaagtga ttggcaccta cctgatgctg 6900
aatttgggag ggtgagatgc atggtgtcga gggaaaggga agaaagacca ggtataggaa 6960
gaagggaaag agcacagcca cagctgtcca gctgtagggt gaagccagtg ggcatagaaa 7020
gtgagggaca gagtcaagag aaaatgcaga tgaatttgat tgggccagag gcggcccagc 7080
agggaggctc tcagacgcac atgaaagcct atggggctgt gggaagccac agccttctct 7140
tccctggcct ggctgtccct gcccactgca cgcatccccc atccgaggtc tgcttctcct 7200
ccctcctagt acaaggagct acagcactct gctagcctgc acgggatgac ctccacacca 7260
ccaagatgga gatctctgag ataaactggt aatgcagagg ctgcactctg agattgataa 7320
cttgtagaag caggtagggc tgagctccca ggtgtccttc catccatgct cctgcctcca 7380
ggtggacttg tgcctgggcc tcactcagcc cctacagaga gaagcctccc agatgagtgg 7440
ccttccagga atataaatat gcattcacct agtgcaatgg ttaatattgt caacttgatt 7500
ggattgaaga atgcgagaaa ggcactttct tctggaagag taaatggcaa aactttggct 7560
tgctttactt actcgcaatc cagaaaagtt cagcattgaa aaggacctca gagggctggg 7620
tgtggtgact cacgcctgta atcccagcac tttgggaggc tgaagagggc ggatcacctg 7680
aggtcaggag ttcgagacca gcctgacgaa catggtgaaa tcccgtctcc actaaaaata 7740
caaaaattag ccaggcatgg tggtgcgcgc ctataatccc agctactcag gaggctgagg 7800
caggggaatc gcttgaacct gggagatgga ggttgcagtg agcagaaatc gtgccagtgc 7860
actccaacct gggcgacaga gtgaaactcc atctcaaaaa aaaaaaaaaa agaaagaaag 7920
aaaaggacct cagaacatca ttttctccat tctggctaag tctcaaacta cccaaggtgg 7980
acagaaatca atcttagttt taaagatgat gcagctccaa cctcctttgc attaaatcca 8040
cttcagggaa ttctttcata ttttcaactt aaagttcata gaggatctta ctgcatcctc 8100
catgaacaat gtcacacaca ttctcatcgc cacctggcag tgtgccatgc cccaggccac 8160
catcgcctgc tgatgccaag cagcatggga agctggcact aagaacaagc tggtggagct 8220
ggagatagcc ctgcagaagg ccaagcagga catggcgcag cagctgtgca ggtaccagga 8280
gccaatgtat gtcaagctag ccccggacat tgagatcgcc acctacagga agctgctgga 8340
gggcttgaag agcaggtggg ccacccctcc cctcccacta ccacccaaac agatgcacca 8400
gaccacgaca gcaatgccac agaggcccca tgagcgccac agtgtcagcc ctccagatga 8460
agctcccagg tcagaaacag ggttaaaatg caattgctcc cttaaaggcg tcctgcagtt 8520
gcagcccctc ccacagctgt tggacccagg gcaggggtga gggtcaaagg ggtcaaggaa 8580
gggacgctaa gcaagaccct ttcttcccaa gtcctgtgag aaatgcaatc atgcttggga 8640
caaggctcag gggctctgga attgccctgc aagctgacgg tctcgcacta cctctctacc 8700
acagctgcct cttctcattt ctctctgtat ttgtgactcg ttctctctcc tggattttcc 8760
tcacagatgc acttgagctt cccaaggcta gcaagctgag gggagagctt gccaagcaca 8820
gtgcctgata catgtaagca ctctgtgctg ataactttaa atacaaagaa cccagggccc 8880
cttaacctct cagcaaacaa aatgactctt tctctcccta gcggtcacag actcactgaa 8940
ggtgtcagag ccgtcagtac ctgtgagtca atgtcttgga atggtagaaa agtgggggc 9000
acaatggaag ctgggagcat cactcttgag ggtctggtgg gtcaaagatg agcaggcagg 9060
ggcagggatg tcacaaagac cctcttggag gctctatggg gactcttctt gtgcccttct 9120
tccccatgag gtcaagagaa ctcaaaaggg gcacattcca tgtcactaag gagctcaaca 9180
actttctcct cactctggct taatctccac aaaaactgtg gcgattaatc tgcaggtgac 9240
ccctcacccc cagggctggt cccagaagat cccagtgaca ggacaagact attcaaacaa 9300
tcgtgggtcc ccagtctgaa tgcctgattt ctcagctagt gctgtcagga actcaaacta 9360
acacaaataa gaggctatta aacagattaa atgtaaaggc ctatatttag aattgaagca 9420
tatgaattca acaaatagaa cagtgactga tctggcgtgg cctcaactca caggagggag 9480
agcagcagtg acctcacaga gtgacccacc agggctctgg ggctgtgtgt aaaagctcag 9540
catcataggc ggcattaatg ggagcattga gtcctggtca tgggagtaga atccccttga 9600
gacccatgtg ggtcagatca cattagagag gctggcctgc caggattggg gagagcctgg 9660
aaccaaccca tcagaggggc aggaatgcct caccaggaaa aacggacgcc aagaggggaa 9720
atcaatgaga cgtttccaaa tgtttacctg gcttgcaaag atgcatccaa gtagggagtt 9780
cccgttgctg aaaatattga agcagaagct gaacaattgt atcttggcaa tgtcattgag 9840
gggttctggt actgggggag gctcgtccag gtggccttca ggcctctttc gaaaaggaag 9900
tctatggctc ccttccagaa tacaaggcct cctcttccat ccaatgtgac agccgcctac 9960
aagacagggg aatgaaccca gcacaggagc taaaagcctt ggcttctggt gggcgtggac 10020
ttgggtatgc cacttgatct ctttgcacct catttttcc tctacaaaat gggggttatt 10080

gtagtaatgg tgtcctctca gccccattga gatccacatg gtcacagctg tgaatgtgcc 10140
ttgagtatgg cgatgccctg tgcagaggtt catgagaagg tagcgtgtga cggggtggtg 10200
agtccacgcc caccctctct cacactggct caggcgttcc ccgaggctct cagggaaggg 10260
agggttgcct ggaacaaggc attgcgggag gggagcaggc atggggatca cacccaggaa 10320
gggcgtcgtg ggaaatggga agaacctgcc aacatatcct gagccctctc tccacagctg 10380
aggtctcttc ccctggggcc acaggctcca gtgggggcag cgtcctctgt gtgggtgggg 10440
gtggccacag cagtggcctc tgccacaaca acaggagcag tggcatcagc tgtccagctg 10500
tgaccagcag tcacagcagc tccaacatat tcattgtctc caagagatgg cccaccaagg 10560
agagtcacag cagccatgca cctgcctcca ctgccactcc cctcaagccc tcaagcttcc 10620
cagctccaag gcttgatatt tactgacgtc ctacacacaa aatcaacagc aggaggtaca 10680
aggagcagga agtggcttat gaggtcatcc tgtccatccc tctgcctcca gaaggtcccc 10740
acagaggtac cccacctgct cccgatctct gcagggtcct gcagggtcca caggcatcca 10800
ttccaaccaa tgttgactca actgctctgt gccaggacct gggctagcac tttgccctta 10860
agaagcttcc agagtagggg gtgcagtagg gtctttcccc tcctgtctaa gtcagaccca 10920
cagcaaccac aagacaaggc ctcttatgca gtccttccca ggacctcaga actgaccacg 10980
gtgccccaaa gagactaagt gcagtgtggg cccaagtttc cacacacctt tctcattgca 11040
gatggaacag ttttcccact caggccagga cacatcttgc acaccctgcc acacatgcca 11100
cctcctcacc tcctttcttc cccctgccca ggccccagcc aagtcccctg acaactccca 11160
ccagcaggat atttgagtga gaaatcaaca cccttcaaaa tatgatcaac tcttgcttat 11220
tcataatata atggaatgat gttatcaaga actcaaaagc cctccatggc ccagggctgc 11280
aatgaatgag agttcatggc caaatttctg tcaagtctgc tttatttaag tcaatattag 11340
acctgctgtt cctaaacaga aatgccctga tgacagagga gtgggaggga agcaggctca 11400
tgtgagggtg tgggaaggcc acttggccca aaatgtttgc aagaaataca ccaggaaaag 11460
cctattttaa ttctcctgca ttaggaactg ccctgtgaga ttcctgcaaa ataatcctcc 11520
aggctcaga 11529

<210> 375

<211> 2218

<212> DNA

<213> Homo sapiens

<400> 375
catggccagc acatccacca ccatcaggag ccacagcagc agccgccggg gtttcagtgc 60
cagctcagcc aggctccctg ggtcagccg ctctggcttc agcagcatct ccgtgtcccg 120
ctccaggggc agtggtggcc tgggtggcgc atgtggagga gctggctttg gcagccgcag 180
tctgtatggc ctgggggggct ccaagaggat ctccattgga gggggcagct gtgccatcag 240
tggcggctat ggcagcagag ccggaggcag ctatggcttt ggtggcgccg ggagtggatt 300
tggtttcggt ggtggagccg gcattggctt tggtctgggt ggtggagccg gccttgctgg 360
tggctttggg ggccctggct tccctgtgtg ccccccctgga ggcatccaag aggtcactgt 420
caaccagagt ctcctgactc ccctcaacct gcaaattgac cccgccatcc agcgggtgcg 480
ggccgaggag cgtgagcaga tcaagaccct caacaacaag tttgcctcct tcatcgacaa 540
ggtgcggttc ctagagcagc agaacaaggt tctggacacc aagtggaccc tgctgcagga 600
gcagggcacc aagactgtga ggcagaacct ggagccgttg ttcgagcagt acatcaacaa 660
cctcaggagg cagctggaca gcatcgtggg ggaacggggt cgtctggact cggagctgag 720
aaacatgcag gacctggtgg aggacctcaa gaacaaatat gaggatgaaa tcaacaagcg 780
cacagcagca gagaatgaat ttgtgactct gaagaaggat gtggatgctg cctacatgaa 840
caaggttgaa ctgcaagcca aggcagacac tcttacagat gagatcaact tcctgagagc 900
cttgtatgat gcagagctgt cccagatgca gacccacatc tcagacacat ccgtggtgct 960
atccatggac aacaaccgca acctggacct ggacagcatc atcgctgagg tcaaggccca 1020
atatgaggag attgctcaga ggagcagggc tgaggctgag tcctggtacc agacaaagta 1080
cgaggagctg caggtcacag caggcagaca tggggacgac ctgcgcaaca ccaagcagga 1140
gattgctgag atcaaccgca tgatccagag gctgagatct gagatcgacc acgtcaagaa 1200
gcagtgtgcc aacctacagg ctgccattgc tgatgctgag cagcgtgggg agatggccct 1260
caaggatgct aagaacaagc tggaagggct ggaggatgcc ctgcagaagg ccaagcagga 1320
cctggcccgg ctgctgaagg agtaccagga gctgatgaac gtcaagctgg ccctggacgt 1380
ggagatcgcc acctaccgca agctgctgga gggcgaggag tgcaggctga atggcgaagg 1440
cgttggacaa gtcaacatct ctgtagtgca gtccaccgtc tccagtggct atggcggtgc 1500
cagcggtgtc ggcagtggct taggcctggg tggaggaagc agctactcct atggcagtgg 1560
tcttggcgtt ggaggcggct ttagttccag cagcggcaga gccactgggg gtggcctcag 1620
ctctgttgga ggcggcagtt ccaccatcaa gtacaccacc acctcctcct ccagcaggaa 1680
gagctacaag cactgaagtc gtgccgccag ctctcagtcc cacagctctc aggcccctct 1740
ctggcagcag agccctctcc tcaggttgct tgtcctcccc tggcctccag tctcccctgc 1800

cctcccgggt agagctggga tgccctcact tttcttctca tcaatacctg ttccactgag 1860
ctcctgttgc ttaccatcaa gtcaacagtt atcagcactc agacatgcga atgtcctttt 1920
tagttcccgt attattacag gtatctgagt ctgccataat tctgagaaga aaaatgacct 1980
atatccccat aagaactgaa actcagtcta ggtccagctg cagatgagga gtcctctctt 2040
taattgctaa ccatcctgcc cattatagct acactcagga gttctcatct gacaagtcag 2100
ttgtcctgat cttctcttgc agtgtccctg aatggcaagt gatgtacctt ctgatgcagt 2160
ctgcattcct gcactgcttt ctctgctctc tttgccttct tttgttctgt tgaataaa 2218

<210> 376

<211> 1986

<212> DNA

<213> Homo sapiens

<400> 376
cgcgccaacg ctcgccacag ccctctcatc tcctggaacc atggccagca catccaccac 60
catcaggagc cacagcagca gccgccgggg tttcagtgcc aactcagcca ggctccctgg 120
ggtcagccgc tctggcttca gcagcatctc cgtgtcccgc tccaggggca gtggtggcct 180
gggtggtgca tgtggaggag ctggctttgg cagccgcagc ttatatggcc tgggggggatc 240
caagaggatc tccattggag ggggcagctg tgccatcagt ggcggctatg gcagcagagc 300
cagaggcagc tatggctttg gtggcgccgg gagtggattt ggtttcggtg gtggagccgg 360
cattggcttt gatctgggtg gtggagccgg ccttgctggt ggctttgggg gccctggctt 420
ccctgtgtgc ccccctggag gcatccaaga ggtcactgtc aaccagagtc tcctgactcc 480
cctcaacctg caaattgacc ccgccatcca gcgggtgcgg gccgaggagc gtgagcagat 540
caagaccctc aacaacaagt ttgcctcctt catcgacaag gtgcggttcc tagagcagca 600
gaacaaggtt ctggacacca agtggaccct gctgcaggag cagggcacca agactgtgag 660
gcagaacctg gagccgttgt tcgagcagta catcaacaac ctcaggaggc agctggacaa 720
catcgtgggg gaacggggcc gcctggactc ggagctgaga aacatgcagg acctggtgga 780
ggacctcaag aacaaatatg aggatgaaat caacaagcgc acagcagcag agaatgaatt 840
tgtgactctg aagaaggatg tggatgctgc ctacatgaac aaggttgaac tgcaagccaa 900
ggcagacact ctcacagatg agatcaactt cctgagagcc ttgtatgatg cagagctgtc 960
ccagatgcag acccacatct cagacacatc cgtggtgcta tccatggaca acaaccgcaa 1020
cctggacctg gacagcatca tcgctgaggt caaggcccaa tacgaggaga ttgctcagag 1080
gagccgggct gaggctgagt cctggtacca gaccaagtac gaggagctgc aggtcacagc 1140
aggcagacat ggggacgacc tgcgcaacac caagcaggag attgctgaga tcaaccgcat 1200
gatccagagg ctgagatctg agatcgacca tgtcaagaag cagtgtgcca gcctgcaggc 1260
tgccattgct gatgctgagc agcgtgggga gatggccctc aaggatgcta agaacaagct 1320
ggaagggctg gaggatgccc tgcagaaggc caagcaggac ctggcccggc tgctgaagga 1380
gtaccaggag ctgatgaatg tcaagctggc cctggacgtg gagatcgcca cctaccgcaa 1440
gctgctggag ggcgaggagt gcaggctgaa tggcgaaggc attggacaag tcaacgtctc 1500
tgtagtacag tccaccatct ccagtggcta tggcggtgcc agtggtgtcg gcagtggctt 1560
aggcctgggt ggaggaagca gctactccta tggcagtggt cttggcattg gaggtggctt 1620
cagttccagc agtggcagag ccattggggg tggcctcagc tctgttggag gcggcagttc 1680
caccatcaag tacaccacca cctcctcctc cagcaggaag agctacaagc actaaagtgc 1740
tgcctccagc tctcggtccc acagtcctca ggcccttctc tggctgcaga gccgtcttct 1800
caggttgcct gtcgtctcct ggcctctagt cttccctgct ctccgaggta gagctgggta 1860
tggatgctta gtgccctcac ttctctctgt ctatacctgc cccatctgag cacccattgc 1920
tcaccatcag atcaaccttt gattttacat cataatgtat tcaccactgg agcttcactt 1980
tgttac 1986

<210> 377

<211> 2222

<212> DNA

<213> Homo sapiens

<400> 377
catggccagc acatccacca ccatcaggag ccacagcagc agccgccggg gtttcagtgc 60
caactcagcc aggctccctg gggtcagccg ctctggcttc agcagcatct ccgtgtcccg 120
ctccaggggc agtggtggcc tgggtggtgc atgtggagga gctggctttg gcagccgcag 180

cttatatggc ctggggggat ccaagaggat ctccattgga gggggcagct gtgccatcag 240
tggcggctat ggcagcagag ccagagcgag ctatggcttt ggtggcgccg ggagtggatt 300
tggtttcggt ggtggagccg gcattggctt tgatctgggt ggtggagccg gccttgctgg 360
tggctttggg ggccctggct tccctgtgtg ccccccctgga ggcatccaag aggtcactgt 420
caaccagagt ctcctgactc ccctcaacct gcaaattgac cccgccatcc agcgggtgcg 480
ggccgaggag cgtgagcaga tcaagaccct caacaacaag tttgcctcct tcatcgacaa 540
ggtgcggttc ctggagcagc agaacaaggt tctggaaaca aagtggaccc tgctgcagga 600
gcagggcacc aagactgtga ggcagaacct ggagccgttg ttcgagcagt acatcaacaa 660
cctcaggagg cagctggaca gcattgtcgg ggaacggggc cgcctggact cagagctcag 720
aggcatgcag gacctggtgg aggacttcaa gaacaaatat gaggatgaaa tcaacaagcg 780
cacagcagca gagaatgaat ttgtgactct gaagaaggat gtggatgctg cctacatgaa 840
caaggttgaa ctgcaagcca aggcagacac tctcacagac gagatcaact tcctgagagc 900
cttgtatgat gcagagctgt cccagatgca gacccacatc tcagacacat ctgtggtgct 960
gtccatggac aacaaccgca acctggacct ggacagcatc atcgctgagg tcaaggccca 1020
atatgaggag attgctcaga gaagccgggc tgaggctgag tcctggtacc agaccaagta 1080
cgaggagctg caggtcacag caggcagaca tggggacgac ctgcgcaaca ccaagcagga 1140
gattgctgag atcaaccgca tgatccagag gctgagatct gagatcgacc acgtcaagaa 1200
gcagtgcgcc aacctgcagg ccgccattgc tgatgctgag cagcgtgggg agatggccct 1260
caaggatgcc aagaacaagc tggaagggct ggaggatgcc ctgcagaagg ccaagcagga 1320
cctggcccgg ctgctgaagg agtaccagga gctgatgaat gtcaagctgg ccctggacgt 1380
ggagatcgcc acctaccgca agctgctgga gggtgaggag tgcaggctga atggcgaagg 1440
cgttggacaa gtcaacatct ctgtggtgca gtccaccgtc tccagtggct atggcggtgc 1500
cagtggtgtc ggcagtggct taggcctggg tggaggaagc agctactcct atggcagtgg 1560
tcttggcgtt ggaggtggct tcagttccag cagtggcaga gccattgggg gtggcctcag 1620
ctctgttgga ggcggcagtt ccaccatcaa gtacaccacc acctcctcct ccagcaggaa 1680
gagctataag cactaagtgc gtctgctagc tctcggtccc acagtcctca ggcccctctc 1740
tggctgcaga gccctctcct caggttgcct gtcctctcct ggcctccagt ctcccctgct 1800
gtcccaggta gagctgggga tgaatgctta gtgccctcac ttcttctctc tctctctata 1860
ccatctgagc acccattgct caccatcaga tcaacctctg attttacatc atgatgtaat 1920
caccactgga gcttcactgt tactaaatta ttaatttctt gcctccagtg ttctatctct 1980
gaggctgagc attataagaa aatgacctct gctccttttc attgcagaaa attgccaggg 2040
gcttatttca gaacaacttc cacttacttt ccactggctc tcaaactctc taacttataa 2100
gtgttgtgaa cccccaccca ggcagtatcc atgaaagcac aagtgactag tcctatgatg 2160
tacaaagcct gtatctctgt gatgatttct gtgctcttca ctctttgcaa ttgctaaata 2220
aa 2222

<210> 378

<211> 2270

<212> DNA

<213> Homo sapiens

<400> 378
ctcctccagc ctctcacact ctcctcagct ctctcatctc ctggaaccat ggccagcaca 60
tccaccacca tcaggagcca cagcagcagc cgccggggtt tcagtgccaa ctcagccagg 120
ctccctgggg tcagccgctc tggcttcagc agcgtctccg tgtcccgctc caggggcagt 180
ggtggcctgg gtggtgcatg tggaggagct ggctttggca gccgcagtct gtatggcctg 240
gggggctcca agaggatctc cattggaggg ggcagctgtg ccatcagtgg cggctatggc 300
agcagagccg gaggcagcta tggctttggt ggcgccggga gtggatttgg tttcggtggt 360
ggagccggca ttggctttgg tctgggtggt ggagccggcc ttgctggtgg ctttgggggc 420
cctggcttcc ctgtgtgccc ccctggaggc atccaagagg tcaccgtcaa ccagagtctc 480
ctgactcccc tcaacctgca aatcgatccc accatccagc gggtgcgggc tgaggagcgt 540
gaacagatca agaccctcaa caacaagttt gcctccttca tcgacaaggt gcggttcctg 600
gagcagcaga acaaggttct ggaaacaaag tggaccctgc tgcaggagca gggcaccaag 660
actgtgaggc agaacctgga gccgttgttc gagcagtaca tcaacaacct caggaggcag 720
ctggacagca ttgtcgggga acggggccgc ctggactcag agctcagagg catgcaggac 780
ctggtggagg acttcaagaa caaatatgag gatgaaatca acaagcgcac agcagcagag 840
aatgaatttg tgactctgaa gaaggatgtg gatgctgcct acatgaacaa ggttgaactg 900
caagccaagg cagacactct cacagacgag atcaacttcc tgagagcctt gtatgatgca 960
gagctgtccc agatgcagac ccacatctca gacacatctg tggtgctgtc catggacaac 1020
aaccgcaacc tggacctgga cagcatcatc gctgaggtca aggcccaata tgaggagatt 1080
gctcagagaa gccgggctga ggctgagtcc tggtaccaga ccaagtacga ggagctgcag 1140
gtcacagcag gcagacatgg ggacgacctg cgcaacacca agcaggagat tgctgagatc 1200

aaccgcatga tccagaggct gagatctgag atcgaccacg tcaagaagca gtgcgccaac 1260
ctgcaggccg ccattgctga tgctgagcag cgtggggaga tggccctcaa ggatgccaag 1320
aacaagctgg aagggctgga ggatgccctg cagaaggcca agcaggacct ggcccggctg 1380
ctgaaggagt accaggagct gatgaatgtc aagctggccc tggacgtgga gatcgccacc 1440
taccgcaagc tgctggaggg tgaggagtgc aggctgaatg gcgaaggcgt tggacaagtc 1500
aacatctctg tggtgcagtc caccgtctcc agtggctatg gcggtgccag tggtgtcggc 1560
agtggcttag gcctgggtgg aggaagcagc tactcctatg gcagtggtct tggcgttgga 1620
ggtggcttca gttccagcag tggcagagcc attgggggtg gcctcagctc tgttggaggc 1680
ggcagttcca ccatcaagta caccaccacc tcctcctcca gcaggaagag ctataagcac 1740
taaagtgcgt ctgctagctc tcggtcccac agtcctcagg cccctctctg gctgcagagc 1800
cctctcctca ggttgcctgt cctctcctgg cctccagtct cccctgctgt cccaggtaga 1860
gctggggatg aatgcttagt gccctcactt cttctctctc tctctatacc atctgagcac 1920
ccattgctca ccatcagatc aacctctgat tttacatcat gatgtaatca ccactggagc 1980
ttcactgtta ctaaattatt aatttcttgc ctccagtgtt ctatctctga ggctgagcat 2040
tataagaaaa tgacctctgc tccttttcat tgcagaaaat tgccaggggc ttatttcaga 2100
acaacttcca cttactttcc actggctctc aaactctcta acttataagt gttgtgaacc 2160
cccacccagg cagtatccat gaaagcacaa gtgactagtc ctatgatgta caaagcctgt 2220
atctctgtga tgatttctgt gctcttcact gtttgcaatt gctaaataaa 2270

<210> 379

<211> 2301

<212> DNA

<213> Homo sapiens

<400> 379
tcgacagctc tctcgcccag cccagttctg gaagggataa aaagggggca tcaccgttcc 60
tgggtaacag agccaccttc tgcgtcctgc tgagctctgt tctctccagc acctcccaac 120
ccactagtgc ctggttctct tgctccacca ggaacaagcc accatgtctc gccagtcaag 180
tgtgtccttc cggagcgggg gcagtcgtag cttcagcacc gcctctgcca tcaccccgtc 240
tgtctcccgc accagcttca cctccgtgtc ccggtccggg ggtggcggtg gtggtggctt 300
cggcagggtc agccttgcgg gtgcttgtgg agtgggtggc tatggcagcc ggagcctcta 360
caacctgggg ggctccaaga ggatatccat cagcactaga ggaggcagct tcaggaaccg 420
gtttggtgct ggtgctggag gcggctatgg ctttggaggt ggtgccggta gtggatttgg 480
tttcggcggt ggagctggtg gtggctttgg gctcggtggc ggagctggct ttggaggtgg 540
cttcggtggc cctggctttc ctgtctgccc tcctggaggt atccaagagg tcactgtcaa 600
ccagagtctc ctgactcccc tcaacctgca aatcgacccc agcatccaga gggtgaggac 660
cgaggagcgc gagcagatca agaccctcaa caataagttt gcctccttca tcgacaaggt 720
gcggttcctg gagcagcaga acaaggttct ggacaccaag tggaccctgc tgcaggagca 780
gggcaccaag actgtgaggc agaacctgga gccgttgttc gagcagtaca tcaacaacct 840
caggaggcag ctggacagca tcgtggggga acggggccgc ctggactcag agctgagaaa 900
catgcaggac ctggtggaag acttcaagaa caagtatgag gatgaaatca acaagcgtac 960
cactgctgag aatgagtttg tgatgctgaa gaaggatgta gatgctgcct acatgaacaa 1020
ggtggagctg gaggccaagg ttgatgcact gatggatgag attaacttca tgaagatgtt 1080
ctttgatgcg gagctgtccc agatgcagac gcatgtctct gacacctcag tggtcctctc 1140
catggacaac aaccgcaacc tggacctgga tagcatcatc gctgaggtca aggcccagta 1200
tgaggagatt gccaaccgca gccggacaga agccgagtcc tggtatcaga ccaagtatga 1260
ggagctgcag cagacagctg gccggcatgg cgatgacctc cgcaacacca agcatgagat 1320
cacagagatg aaccggatga tccagaggct gagagccgag attgacaatg tcaagaaaca 1380
gtgcgccaat ctgcagaacg ccattgcgga tgccgagcag cgtggggagc tggccctcaa 1440
ggatgccagg aacaagctgg ccgagctgga ggaggccctg cagaaggcca agcaggacat 1500
ggcccggctg ctgcgtgagt accaggagct catgaacacc aagctggccc tggacgtgga 1560
gatcgccact taccgcaagc tgctggaggg cgaggaatgc agactcagtg gagaaggagt 1620
tggaccagtc aacatctctg ttgtcacaag cagtgtttcc tctggatatg gcagtggcag 1680
tggctatggc ggtggcctcg gtggaggtct tggcggcggc ctcggtggag gtcttgccgg 1740
aggtagcagt ggaagctact actccagcag cagtgggggt gtcggcctag gtggtgggct 1800
cagtgtgggg ggctctggct tcagtgcaag cagtggccga gggctggggg tgggctttgg 1860
cagtggcggg ggtagcagct ccagcgtcaa atttgtctcc accacctcct cctcccggaa 1920
gagcttcaag agctaagaac ctgctgcaag tcactgcctt ccaagtgcag caacccagcc 1980
catggagatt gcctcttcta ggcagttgct caagccatgt tttatccttt tctggagagt 2040
agtctagacc aagccaattg cagaaccaca ttctttggtt cccaggagag ccccattccc 2100
agcccctggt ctcccgtgcc gcagttctat attctgcttc aaatcagcct tcaggtttcc 2160
cacagcatgg cccctgctga cacgagaacc caaagttttc ccaaatctaa atcatcaaaa 2220

cagaatcccc accccaatcc caaattttgt tttggttcta actacctcca gaatgtgttc 2280
aataaaatgc ttttataata t 2301
<210> 380

<211> 2255

<212> DNA

<213> Homo sapiens

<400> 380
ccttcctccc tcctgcatct gagctttgtc tccaccagca acatgagccg ccaattcacc 60
tgcaagtcgg gagctgccgc caaggggggc ttcagtggct gctcagctgt gctctcaggg 120
ggcagctcat cctccttccg ggcagggagc aaagggctca gtgggggctt tggcagccgg 180
agcctctaca gcctgggggg tgtccggagc ctcaatgtgg ccagtggcag cgggaagagt 240
ggaggctatg gatttggccg ggccgggcc agtggctttg ctggaagcat gtttggcagt 300
gtggccctgg ggcctgtgtg cccaactgta tgcccacctg gaggcatcca ccaggttacc 360
gtcaatgaga gcctcctggc cccccctcaac gtggagctgg accccgagat ccagaaagtg 420
cgtgcccagg agcgagagca gatcaaggct ctgaacaaca agttcgcctc cttcatcgac 480
aaggtgcggt tcctggagca gcagaaccag gtactggaga ccaagtggga gctgctgcag 540
cagctggacc tgaacaactg caagaacaac ctggagccca tcctcgaggg ctacatcagc 600
aacctgcgga agcagctgga gacgctgtct ggggacaggg tgaggctgga ctcggagctg 660
aggaatgtgc gggacgtagt ggaggactac aagaagaggt atgaggagga aatcaacaag 720
cggacagcag cagagaacga gtttgtgctg ctcaagaagg atgtggatgc tgcttacgcc 780
aataaggtgg aactgcaggc caaggtggaa tccatggacc aggagatcaa gttcttcagg 840
tgtctctttg aagccgagat cactcagatc cagtcccaca tcagtgacat gtctgtcatc 900
ctgtccatgg acaacaaccg gaacctagac ctggacagca tcattgacga agtccgcacc 960
cagtatgagg agattgcctt gaagagtaag gccgaggctg aggccctgta ccagaccaag 1020
ttccaagagc ttcagctggc agctggcagg catggggacg acctcaaaaa caccaagaat 1080
gaaatctcgg agctcactcg gctcatccag agaatccgct cagagatcga gaacgtgaag 1140
aagcaggctt ccaacctgga gacagccatc gctgatgctg agcagcgggg agacaacgcc 1200
ctgaaggatg cccgggccaa gctggacgag ctggagggcg ccctgcacca ggccaaggag 1260
gagctggcgc ggatgctgcg cgagtaccag gagctcatga gcctgaagct ggccctggac 1320
atggagatcg ccacctatcg caagctactg gagagcgagg agtgcaggat gtcaggagaa 1380
tttccctccc ctgtcagcat ctccatcatc agcagcacca gtggcggcag tgtctatggc 1440
ttccggccca gcatggtcag cggtggctat gtggccaaca gcagcaactg catctctgga 1500
gtgtgcagcg tgagaggcgg ggagggcagg agccggggca gtgccaacga ttacaaagac 1560
accctaggga agggttccag cctgagtgca ccctccaaga aaaccagtcg gtagagaaga 1620
ctgccccggg ccccgcctca ttccatgacc cggctctgga tcccacactg tacttcccac 1680
agcccactct cagctccatc tccaccctgc tggtcctgct cccatacacc tggcactggc 1740
cttggccacc cacttctccc agcctgtgtc ttcctgatcc tgggaaggcc tggatgacca 1800
agcttggtga aattcctccc tgtacacacc ctattaactc cttggctgtg gtcccccagc 1860
tacaccacca gcccaggtcc tggctgccag ctttcctcct ctgcccggcc tctagcgcag 1920
tcgctaacta ctctgctggg ctccctgggt ctctgcccaa ggccccgcac acactgggac 1980
ctagcatagt tcctgcctat gccaggagct ggctctgtgt ttaagaaaag gaggactgaa 2040
ggacaaacaa ccaagagtgg cccagtcccc acccccacat ctagctcagt ctcaaatctg 2100
agtgggacca agtgcaattc agggcctttt tctccactca cctgcaccca gaagcagaga 2160
aaagcaggca ctgttcactt ttcctttatt cttaatggcc ttcctctgtt gcaacctcaa 2220
taaacagcac aatctcaaaa aaaaaaaaaa aaaaa 2255
<210> 381

<211> 2856

<212> DNA

<213> Homo sapiens

<400> 381
ccttggagac tgcttttctc cagctctgtc aactcaacct ttcccaccat gagtcggcaa 60
ctgaacatca agtccagtgg tgacaagggc aacttcagtg tgcattcggc agtggtgcca 120
aggaaggctg tgggtagcct ggcttcttac tgtgcagctg gcagaggggc tggcgctggc 180
tttggcagtc ggagcctcta tagccttgga gggaatcggc gtatttcttt caatgtggct 240

ggtggcggcg ttcgggctgg aggttacggc ttcaggcctg gctctgggta tggagggggc 300
cgggccagtg gctttgctgg cagtatgttt ggcagtgtgg ccctggggcc tgcatgtttg 360
tctgtgtgcc cacctggggg catccaccag gtcactgtca acaagagcct cttggccccc 420
ctcaacgtgg agctggaccc tgagatccag aaggtgcgcg cccaggagcg ggaacagatc 480
aaggtgctga acgacaagtt cgcctccttc attgacaagg tacgcttcct agagcagcag 540
aaccaggttc tagaaaccaa gtgggagctg ctgcagcagc tggacctgaa caactgcaag 600
aagaacctgg agcccatcct tgagggctac atcagcaacc tgcggaagca gctggagaca 660
ctgtctgggg acagggtgag gctggactcg gagctgagaa gcatgaggga tctggtggag 720
gactataaga agagatatga ggtggagatt aaccggcgca caacagcaga gaatgagttt 780
gtggtgctta agaaggatgc agatgcagcc tacgcagtca aggtggagct tcaggccaaa 840
gtggactcac tggacaaaga catcaagttc ctcaagtgtc tgtatgatgc agagatcgct 900
cagatccaga ctcacgccag tgagacctct gtcatcctgt ccatggacaa caaccgggac 960
ctggaccttg acagcatcat cgctgaggtc cgcatgcatt atgaggagat cgccctgaag 1020
agcaaggccg aggccgaggc cctgtaccag accaagatcc aggagctgca gctggcagcc 1080
agtcggcatg gtgacgacct gaaacacacc aggagcgaga tggtggagct gaaccggctc 1140
atccagagga tccggtgtga gatcgggaat gtgaagaagc agcgtgccag cctggagacg 1200
gccatcgctg acgctgagca gcggggagac aatgccctga aggatgccca ggccaagctg 1260
gatgagctgg agggcgccct gcaccaggcc aaggaggagc tggcgcggat gctgcgcgag 1320
taccaggagc tcatgagcct gaaactggcc ctggacatgg agattgccac ctaccgcaag 1380
ctgctggagg gcgaggagtg caggatgtct ggtgagaatc catcctctgt gagcatctct 1440
gtcatcagca gtagcagcta cagctaccac caccccagct ctgcgggtgt tgaccttggg 1500
gccagcgctg tggcaggcag ctctggcagc acccagagcg ggcagaccaa gaccacagag 1560
gcgcgagggg gagacctcaa ggacacccag ggcaagagca ccccagccag catcccagca 1620
aggaaagcca cccgctagac ccatggcctc acccacctca gcacttggaa gaagaggtga 1680
ctttgccacc cccaaaggtg tctgccacac ccaagttccc aggccctgag ttttaaaact 1740
gtctgtagta cactcaactg tctgcatcgt ggtttagctt ttactttcaa gctctgattg 1800
acacagtcac cttccctgtt tccttaggtc ccatgtggac taacgacttc tcattttcct 1860
cgctgccttt ggctggcagg aggctttgga ggcacaagcc attataacct tcttggccct 1920
aaggaagctg tgatcatccc tagaaagagg gaaggagcag gagacgacag gggaggggct 1980
ggtttgttct gtgcttaggc caagttagct actgtctgag aggttttaca tcccctgcca 2040
gcatggtggg gtgacacgag acctgtaagc aggtggtgag aagttagaca gcctttactt 2100
gctccatcag aaacaacttt gcaggatgac gccaatattt aggacaccac gtgtacatct 2160
gtagaccacc agctgcaccc atctccacaa agccaagtga aggtgtattg gggattctct 2220
gccagcctgt gtcacccacc cacctccttc attatcagca atagctacca tcatttgtca 2280
agcacctact agatgccaga caccctacac atattgcctc ctattctcat ccattcctgc 2340
aaagaagatg caatgagcat ccttggaaca gactgggagc ctgatcccaa gaggctgctg 2400
aactggcccc agtcacccac ctatgtgcca gagtcaggcc tggccgacac aaggatctat 2460
gctttttctg tggtgtgtca ctgccttttt aacaaaaggg ctgtcaaagt caccattctt 2520
tttgatgagg gcagcgtaat tacattgcct tttgtcagga ctgtgggtat gtgatttgtc 2580
ttccacagtc tctctggggc tgtgtcctac atagcttctg actctcaatt tttggtgcca 2640
tgagctgagc tcagtgagtg gagctggctt ttctgtcgct gagtggtact gcctagtcca 2700
gatgtcgtgg ctcagggtgt aaatattcct aagaatggca gcatctattt cctcttttgt 2760
ttgaattaaa gactctgaat ttctgctgaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2820
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 2856

<210> 382

<211> 1985

<212> DNA

<213> Homo sapiens

<400> 382

gcggctgttt gcgccccttc tgcctcccgc gcggccagca ccggtacttg cgagccatga 60
gccgccaact gacccatttc ccccgcgggg agcgcctggg cttcagcggt tgctccgcgg 120
tcctctctgg cgggatcggc agcagctccg cctcattccg ggcccgggtc aagggctcgg 180
cctcctttgg cagcaagagc ctctcctgcc ttgggggcag ccgaagcctg gcgctcagcg 240
ctgctgcacg gcgggggcggc ggccgcctgg gcggcttcgt gggcaccgcc ttcggcagcg 300
ccgggctggg gcccaagtgt ccctccgtgt gcccaccccgg gggcatccct caggtcaccg 360
tcaacaagag cctcctggcc ccgctcaacg tggagatgga ccccgagatc cagagggtgc 420
gcgcccagga gcgggagcag atcaaggcgc taaacaacaa gttcgcctcc ttcatcgaca 480
aggtgcggtt cctggagcag cagaatcagg tgctagagac caagtggaac ctcctacagc 540
agctggactt gaacaactgc aggaagaacc tggagcccat ttatgagggc tacatcagca 600
acctgcagaa gcagctggag atgctgtctg ggacggggt gaggctggat tcggagctga 660

ggaacatgca ggatttggtg gaggactaca agaagaggta tgaggtggag attaacagac 720
gcacagctgc tgagaatgag tttgtggtgc tcaagaagga cgtggatgct gcttacatga 780
ataaggttga gctccaggcc aaggtggact ccttgacaga tgagattaaa ttcttcaagt 840
gcctttatga aggggagatc actcagatcc agtcccacat cagcgacacg tccatcgtcc 900
tgtcaatgga caacaaccgg gatctggacc tggacagcat cattgccgag gtccgtgccc 960
agtacgagga gattgcccta aagagcaagg ccgaggctga gaccctgtac cagaccaaga 1020
tccaggagct gcaggtcaca gcaggccagc atggggatga cctcaagctc accaaggctg 1080
aaatctctga gctcaaccgc ctgatccaga ggatccgctc agagataggg aatgtgaaga 1140
agcagtgtgc cgatctggag acggccatcg ccgacgctga acagcggggg gactgcgccc 1200
tgaaagatgc ccgggccaag ctggatgagc tggagggcgc cctgcaccag gccaaggagg 1260
agctggcacg gatgctgcgt gagtaccagg agctcgtgag cctgaagctg gccctggata 1320
tggagatcgc cacctaccgc aagctgctgg agagcgagga gtgcaggatg tctggcgaat 1380
atccaaattc tgtgagcatc tccgtcatca gcagcaccaa tgctggggca ggaggggctg 1440
gcttcagcat gggctttggc gcctcaagca gttatagcta caaaactgca gctgcagacg 1500
tcaagaccaa aggcagctgt ggcagtgagc tcaaggatcc ccttgccaaa acctcgggga 1560
gcagctgtgc caccaaaaag gcctccagat gatggacaag tggttggctc tgtgagcaga 1620
aggcttccca actcacccgt ctcctccctc tccttccctg ggctcccttt ccaagtcaag 1680
aaatgctttg tctgtcacta cagtcccaac cccatttcct ctccctgttg tcctcaaggt 1740
gtgatgcttc gatgcaagga gtccagtcag gagctctctc agtggcctcc ttcttggtcc 1800
cgtttcgttg ggtcattgcc tgaaatgcaa agacaacact tacccaggat gtttccccaa 1860
ggccagctac aagcccctct gcccccatgc ctttcttggg ccttcactac caagggactc 1920
tcttcatctt ctgattggga ttgtgtccag tcctctgctt cttctgcaat aaatgattaa 1980
atctg 1985

<210> 383

<211> 2323

<212> DNA

<213> Homo sapiens

<400> 383
tcagcatctt atccccactt tctggcctcc ccaccatgag ccgccaattc acctacaagt 60
cgggagctgc tgccaagggg ggcttcagcg gctgctccgc tgtgctctca gggggcagct 120
catcctccta ccgagcaggg ggcaaagggc tcagtggagg cttcagcagt cggagccttt 180
acagcctggg gggtgcccgg agcatctctt tcaatgtggc cagtggcagt gggtgggcag 240
gaggctatgg atttggccgg ggccgggcca gtggctttgc tggcagcatg tttggcagtg 300
tggccttggg gtccgtgtgt ccgtcgttgt gcccgcccgg gggtatccat caggtcacca 360
tcaacaagag cctcctggca cccctgaacg tggagctgga ccctgaaatc cagaaagtgc 420
gtgcccagga gcgggagcag atcaaggtgc tgaacaacaa gttcgcctcc ttcattgaca 480
aggtgcggtt cctggagcag cagaaccagg tgctggagac caagtgggag ctgctacagc 540
agctggacct gaacaactgc aagaataacc tggagcccat ccttgagggc tacatcagca 600
acctgcggaa gcagctggag acgctgtctg ggacagggt gaggctggac tcggagctga 660
ggagcgtgcg cgaagtggtg gaggactaca agaagaggta tgaagaagaa ataaacaagc 720
gcacaactgc tgagaatgaa tttgtggtgc ttaagaagga cgtggacgca gcttacacga 780
gcaaagtgga gctgcaggcc aaggtggatg ccctggatgg agaaatcaag ttcttcaagt 840
gtctgtacga gggggagact gctcagatcc agtcccacat cagcgacacg tccatcatcc 900
tgtccatgga caacaaccgg aacctggacc tggacagcat cattgctgag gtccgtgccc 960
agtatgagga gatcgcccgg aagagcaagg ccgaggccga ggccctgtac cagaccaagt 1020
tccaggagct gcagctagca gccggccggc atggggatga cctgaaacac accaaaaatg 1080
agatctcaga gctgacccgt ctcatccaaa gactgcgctc ggagattgag agtgtgaaga 1140
agcagtgtgc caacctggag acggccatcg ctgacgccga gcagcggggg gactgtgccc 1200
tcaaggatgc cagggccaag ctggatgagc tggagggcgc cctgcagcag gccaaggagg 1260
agctggcacg gatgctgcgc gagtaccaag agcttttgag cgtgaagctg tccctggata 1320
ttgagatcgc cacctaccgc aagctgctgg agggcgagga gtgcaggatg tccggagaat 1380
ataccaactc cgtgagcatt tcggtcatca acagctccat ggccgggatg gcaggcacag 1440
gggctggctt tggattcagc aatgctggca cctacggcta ctggcccagc tctgtcagcg 1500
ggggctacag catgctgcct gggggctgtg tcactggcag tgggaactgt agcccccgtg 1560
gggaagccag gaccaggctg gggagtgcaa gtgaattcag ggactcccag ggaaagacct 1620
tagctctaag ctcacccacc aaaaaaacca tgagataaaa gtgcaacaca gcccatttcc 1680
cagtagtctg ccttgccctg gcagactgct ccggactctt ctaggaaatt cctctgtctc 1740
atttgtcctc atctccttct tgctttgctt tgctgtggac ctctcctcct tctccttgac 1800
ctctctgcct cagtctttgt taatgttgaa tcaggatttg agagctggag cccatctttc 1860
tgcctccatc caaagaggcc ttttaactca gcctcccctt ctggaagaag acttggcctt 1920

ttctctgcct gcatgtccct gtgcagagga aggctgtctc tcaagggtaa ccccaatctg 1980
cctgtgctgc tctttcatac cctgtctcag taactatttg gccatgtgct ccctgggtac 2040
acaagttttc agatctactg atagacctgc aggaactctc tcatgttaga gcaatactgg 2100
ctcaggcaaa acacctggtg ggaagttgta tcttgtatag ccactgcata ttacatagcc 2160
aggagtaccc cttgaggcta attctaccaa ccaagggatt tgtgctcttc atctgtgggt 2220
acttgtcctt gtttaggcct cacctccctg ctcctgtgtc ttaccaataa acttataaag 2280
cccaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 2323

<210> 384

<211> 2427

<212> DNA

<213> Homo sapiens

<400> 384

agcctgtgac tttcctccct ggacaaaggc atcatgagtt gtcagatctc ttgcaaatct 60
cgaggaagag gaggaggtgg aggaggattc cggggcttca gcagcggctc agctgtggtg 120
tctggtggaa gccggagatc aacttccagc ttctcctgct tgagccgcca tggtggtggt 180
ggcgggggct tcggtggagg cggctttggc agtcggagtc ttgttggcct tggagggacc 240
aagagcatct ccattagtgt ggctggagga ggtggtggct ttggcgccgc tggtggattt 300
ggtggcagag gaggtggttt tggaggcggc agcggctttg gaggcggcag cggctttgga 360
ggtggcagcg gcttcagtgg tggtggtttc ggtggaggcg gctttggtgg aggccgcttt 420
ggaggttttg gggccctggg tggtgttgga ggtttagggg gtcctggtgg ctttgggcct 480
ggaggatacc ctggtggcat ccacgaagtc tctgtcaacc agagcctcct gcagcctctc 540
aacgtgaaag ttgacccaga gatccagaat gtgaaggccc aagagcgtga gcagatcaaa 600
actctcaaca acaaatttgc ctccttcatt gacaaggtgc ggttcttgga gcagcagaac 660
caggtgttac agaccaaatg ggagctgcta caacaaatga atgttggcac ccgccccatc 720
aacctggagc ccatcttcca ggggtatatc gacagcctca agagatatct ggatgggctc 780
actgcagaaa gaacatcaca gaattcagag ctgaataaca tgcaggatct tgtggaggat 840
tataagaaga agtatgagga tgaaatcaat aagcgcacag ctgctgagaa tgatttgtg 900
acgcttaaaa aggacgtgga caatgcctac atgataaagg tggagttgca gtccaaggtg 960
gacctgctga accaggaaat tgagtttctg aaagttctct atgatgcgga gatatcccag 1020
atacatcaga gtgtcactga caccaacgtc atcctctcca tggacaacag ccgcaacctg 1080
gacttggata gcatcatcgc cgaggtcaag gcccagtatg aggagatcgc ccagaggagc 1140
aaggaagaag cggaggccct gtaccacagc aagtatgagg agctccaggt gactgtcggg 1200
agacatggag acagcctgaa agagatcaag atagagatca gcgagctgaa ccgcgtgatc 1260
cagaggctgc aggggggagat cgcacatgtg aagaagcagt gtaagaatgt gcaagatgcc 1320
atcgcagatg ccgagcagcg tggggagcat gccctcaagg atgccaggaa caagttgaat 1380
gacctggagg aggccctgca gcaggccaag gaggacttgg cgcggctgct gcgtgactac 1440
caggagctga tgaacgtgaa gctggcccta gatgtggaga tcgccaccta ccgcaaactg 1500
ctggagggcg aggagtgcag gatgtctgga gacctcagca gcaatgtgac tgtgtctgtg 1560
acaagcagca ccatttcatc aaatgtggca tccaaggctg cctttggagg ttctggaggt 1620
agagggtcca gttccggagg aggatacagc tctggaagca gcagttatgg ctctggaggc 1680
cgacagtctg gctccagagg cggtagtgga ggaggaggtt ctatctctgg aggaggatat 1740
ggctctggcg gtggttctgg aggaagatac ggatctggtg gtggctctaa gggagggtcc 1800
atctctggag gaggatatgg ctctggaggt ggaaaacaca gctctggagg tggctctaga 1860
ggaggctcca gctctggagg aggatatggc tctggaggtg gggttctag ctctgtaaag 1920
ggtagctcag gtgaagcttt tggttccagc gtgaccttct cttttagata aagatgagcc 1980
cccaccacca ccgactctcc caacccagac tctcccactc cagaatgtag aagcctgtct 2040
ctgtacctct aactggcagc aagttaaatt tttgtcattt atctctgatg gcactttgag 2100
ggaaaagaat gtccacatac agtttttgaa agatcttctc tccaaaccag ttagttagag 2160
ccagtgacgc ctctgtgttc tggggcggaa tctgtgctgt ctaggtttgt gcttctagcc 2220
atgcccattc ccgcccccac catgcctctt tgcattgccc attttccaga tgtgtattct 2280
gttgaggacc caggcccatc cagggatttc atctctaagc ctggcagtgc tggggggaaa 2340
tgtgtttctg tgtatatagc tcctcttgtc cactctgctt tcggaagtgc tgtggtctgg 2400
gggtcttcat aataaacctc atttgca 2427

<210> 385

<211> 2507

<212> DNA

<213> Homo sapiens

<400> 385
agctccttcc cttactctac cttgctccta cttttctcta agtcaacatg agtcgacagt 60
ttagttccag gtctgggtac cgaagtggag gggcttcag ctctggctct gctgggatca 120
tcaactacca gcgcaggacc accagcagct ccacacgccg cagtggagga ggtggtggga 180
gattttcaag ctgtggtggt ggtggtggta gctttggtgc tggtggtgga tttggaagtc 240
ggagtcttgt taaccttggt ggcagtaaaa gcatctccat aagtgtggct agaggaggtg 300
gacgtggtag tggctttggt ggtggttatg gtggtggtgg ctttggtggt ggtggctttg 360
gtggtggtgg ctttggtgga ggtggcattg gggtggtgg ctttggtggt tttggcagtg 420
gtggtggtgg ttttggtgga ggtggctttg gggtggtgg atatgggggt ggttatggtc 480
ctgtctgccc tcctggtggc atacaagaag tcactatcaa ccagagcctt cttcagcccc 540
tcaatgtgga gattgaccct gagatccaaa aggtgaagtc tcgagaaagg gagcaaatca 600
agtcactcaa caaccaattt gcctccttca ttgacaaggt gaggttcctg gagcagcaga 660
accaggtact gcaaacaaaa tgggagctgc tgcagcaggt agatacctcc actagaaccc 720
ataatttaga gccctacttt gagtcattca tcaacaatct ccgaaggaga gtggaccaac 780
tgaagagtga tcaatctcgg ttggattcgg aactgaagaa catgcaggac atggtggagg 840
attaccggaa caagtatgag gatgaaatca acaagcggac aaatgcagag aatgaatttg 900
tgaccatcaa gaaggatgtg gatggtgctt atatgaccaa ggtggacctt caggccaaac 960
ttgacaactt gcagcaggaa attgatttcc ttacagcact ctaccaagca gagttgtctc 1020
agatgcagac tcaaatcagt gaaactaatg tcatcctctc tatggacaac aaccgcagtc 1080
tcgacctgga cagcatcatt gctgaggtca aggcccagta cgaggatata gcccagaaga 1140
gcaaagctga ggccgagtcc ttgtaccaga gcaagtatga agagctgcag atcactgctg 1200
gcagacatgg ggatagtgtg agaaattcaa agatagaaat ttctgagctg aatcgtgtga 1260
tccagagact tagatctgaa atcgacaatg tcaagaagca gatctccaac ttgcagcagt 1320
ccatcagtga tgcagagcag cgtggcgaga atgccctcaa ggatgccaag aacaagctga 1380
atgacctgga ggatgccctg cagcaggcca aggaagacct ggcccgcctg ctgcgcgact 1440
accaggagct gatgaacacc aagctggccc tggatctgga gattgccacc tacaggaccc 1500
tcctggaggg agaagaaagc aggatgtctg gagaatgtgc cccgaacgtg agtgtgtctg 1560
tgagcacaag ccacaccacc atcagtggag gtggcagccg aggaggtggc ggcggtggct 1620
acggctctgg aggtagcagc tatggctccg gaggtggtag ctatggttct ggaggtggcg 1680
gcggcggcgg ccgtggcagc tatggctccg gaggtagcag ctacggctcc ggaggtggca 1740
gctatggctc tggaggtggc ggcggcggcc atggcagcta cggctccgga agcagcagtg 1800
gggctacag aggtggctct ggaggcggcg gcggcggcag ctctggcggc cggggctctg 1860
gcggcgggag ctctggaggc tccataggag gccggggatc cagctctggg ggtgtcaagt 1920
cctctggtgg cagttccagc gtgaggtttg tttctaccac ttattccgga gtaaccagat 1980
aaagagatgc cctctgtttc attagctcta gttctccccc agcatcacta acaaatatgc 2040
ttggcaagac cgaggtcgat ttgtcccagc cttaccggag aaaagagcta tggttagtta 2100
cactagctca tcctattccc ccagctcttt cttttctgct gtttcccaat gaagttttca 2160
gatcagtggc aatctcagtc ccctggctat gaccctgctt tgttctttcc ctgagaaaca 2220
gttcagcagt gaccaccacc cacatgacat ttcaaagcac ctccttaagc cagccagagt 2280
aggaccagtt agacccaggg tgtggacagc tccttagcat cttatctctg tgctgttttg 2340
gttttgtaca taaggtgtaa gcaagttgtt tttcttttgt ggagaggtct taaactcccc 2400
atttccttgt tttgctgcaa taaactgcat ttgaaattct ccatgtctcg atcgcccttg 2460
tttacggctc tgtctaacct ggatgggtgt tttgtgaggt aaaagaa 2507

<210> 386

<211> 2224

<212> DNA

<213> Homo sapiens

<400> 386
atgagcagac aagccagcaa gacatctggt ggcgggagcc agggtttctc cgggcgctct 60
gctgtggtct ccggcagcag caggatgagc tgtgtggccc actctggggg agctggcgga 120
ggggcctatg gcttccggag cggagcaggt ggctttggca gtcgcagcct ctacaacctg 180
ggcggcaaca agagcatctc catcagcgtg gcagctggcg gctcccgggc tggaggcttt 240
gggggagggc ggagcagctg tgcctttgca ggtggctatg gaggtggctt tgggagcggc 300
tatggaggtg gctttggtgg tggctttggt ggtggcagag gaatgggagg tggctttggt 360
ggagctggtg gctttggagg ggctggtggc tttggagggg ctggtggctt tggtggtcct 420
ggtggctttg gtgggtctgg tggctttggt gggcctggca gcttgggcag tcctggtggc 480
tttgcgcctg gggctttcc tgggggaatt caggaagtga ctactaacca gagtctcctg 540

cagcccctca aagtggagac tgacccccag attgggcaag taaaggccca ggagcgggaa 600
cagatcaaga ccctcaacaa caagtttgcc tccttcattg acaaggtgcg gttcctggag 660
caacagaaca aagtcctgga gaccaagtgg aacctgctcc agcagcaggg cacaagttcc 720
atctcaggca caaacaacct tgagcctctt tttgagaatc acatcaacta cctgcggagc 780
tacctggaca acatcctcgg ggagagaggg cgcctggact ctgagctgaa gaacatggag 840
gacctggtgg aagacttcaa gaagaaatat gaggatgaaa tcaataaacg tacagctgct 900
gagaatgaat ttgtgactct gaagaaggat gtggacagtg cctatatgaa caaggtggag 960
cttcaggcca aagtggatgc cttgatagat gagatcgact tcttaaggac cctctacgac 1020
gctgagctat ctcagatgca gagccacatc agtgacacat ctgtggtgct gtccatggac 1080
aataatcgtt ccctggacct ggacagcatc attgctgaag ttggtgcaca gtatgaggat 1140
atcgctcaga gaagcaaggc cgaagctgag gccctgtacc agaccaagtt gggggagctg 1200
cagaccacgg ctggcaggca tggggatgac ctaagaaata ccaagagcga gatcatagag 1260
ctcaacagaa tgatccagag gctgcgggca gagatcgagg gtgtcaagaa gcagaatgcc 1320
aacctgcaga cggccattgc gcaggccgag cagcatggag agatggccct caaggatgcc 1380
aatgccaagc tccaagagct gcaggctgct ctacagcagg cgaaggatga cctggcgcgg 1440
ctgctacgtg actaccagga gctgatgaat gtcaagctgg ccctggacgt ggagatcgcc 1500
acctaccgca agctgctgga gggcgaggag tacagcagga tgtctggaga gtgtccgagt 1560
gctgtcagca tctccgtggt cagcagcagc acgacttccg cctccgcagg tggctatgga 1620
ggaggttacg gcggaggcat gggcggtggt ttaggaggtg gcttcagtgc gggcggcggc 1680
tcaggcattg gctttggccg gggaggcggc ggtggaatcg gcggtggatt tggcggcggc 1740
acgagcggtt tcagcggtgg cagcggcttt ggctccatct ctggcgcccg ctatggagtg 1800
tcccagtcct cccagcgcta ctccagataa cggggcggca gcatcaagtt ctcccagtcc 1860
gcgcctcccc tccgctcccc acgcccccat attggcaaca cccccagcgc caccccatgc 1920
tccccaagaa cgcggctctg cgctggaagg ctgtggcctc ttgctcctcc ttccagatcc 1980
ccggccaagc ccagtcacga gaatttcagg ctcttctctc cggctccagc ctgtgattgt 2040
ccaggtgtgg cctaggcgcc accttctgct gccccggagt cccctccctg ggtgggatgc 2100
tcactacagc tctgatgtat atagcacaca gccccctccg cccctccggt ctgtcgccaa 2160
taaa 2220

2224

<210> 387

<211> 2617

<212> DNA

<213> Homo sapiens

<400> 387

gacttgctcc ggtttgcaga gctaggaggt ggcaggctgt gcgctcaaac tcaggctgtc 60
taactccaca ttctgtgggg tgagaggatg ggtgatgggg tgtcttttct ggaggaggga 120
ggtgctgtga gcctagcgag atggaggtac agtgggtgtg ggcctggagc gctgggccca 180
ggcaggggct tctgattagg aagccctggg gcaccagttc aggttctccc agagagtagt 240
gtgatgggat ccagtaacct gtgccctcca gatgacttct gtaggtgtgt ttagtgacat 300
gctcaacggg tgcgggaagg atgggcttgt gccaagggcc aagcccagag atgtttcaga 360
tttttccctt tatgcccctg caaccaagcc ctgctgctcc aggacatata agagacgaag 420
gctgagggct ccagcactca ccggcctggg ccctgtcact tctctgatag ctcccagctc 480
gctctctgca gccatgattg ccagacagca gtgtgtccga ggcgggcccc ggggcttcag 540
ctgtggctcg gccattgtag gcggtggcaa gagaggtgcc ttcagctcag tctccatgtc 600
tggaggtgct ggccgatgct cttctggggg atttggcagc agaagcctct acaacctcag 660
ggggaacaaa agcatctcca tgagtgtggc tgggtcacga caaggtgcct gctttggggg 720
tgctggaggc tttggcactg gtggctttgg tgccggcggc ttcggagctg gtttcggcac 780
tggtggcttt ggtggtggat ttggggggctc cttcagtggt aagggtggcc ctggcttccc 840
cgtctgcccc gctgggggaa ttcaggaggt caccatcaac cagagcttgc tcacccccct 900
ccacgtggag attgaccctg agatccagaa agtccggacg gaagagcgcg aacagatcaa 960
gctcctcaac aacaagtttg cctccttcat cgacaaggtg cagttcttag agcaacagaa 1020
taaggtcctg gagaccaaat ggaacctgct ccagcagcag acgaccacca cctccagcaa 1080
aaaccttgag cccctctttg agacctacct cagtgtcctg aggaagcagc tagatacctt 1140
gggcaatgac aaagggcgcc tgcagtctga gctgaagacc atgcaggaca gcgtggagga 1200
cttcaagact aagtatgaag aggagatcaa caaacgcaca gcagccgaga atgactttgt 1260
ggtcctaaag aaggacgtgg atgctgccta cctgaacaag gtggagttgg aggccaaggt 1320
ggacagtctt aatgacgaga tcaacttcct gaaggtcctc tatgatgcgg agctgtccca 1380
gatgcagacc catgtcagcg acacgtccgt ggtcctttcc atggacaaca accgcaacct 1440
ggacctggac agcattattg ccgaggtccg tgcccagtac gaggagattg cccagaggag 1500
caaggctgag gctgaagccc tgtaccagac caaggtccag cagctccaga tctcggttga 1560

ccaacatggt gacaacctga agaacaccaa gagtgaaatt gcagagctca acaggatgat 1620
ccagaggctg cgggcagaga tcgagaacat caagaagcag tgccagactc ttcaggtatc 1680
cgtggctgat gcagagcagc gaggtgagaa tgcccttaaa gatgcccaca gcaagcgcgt 1740
agagctggag gctgccctgc agcaggccaa ggaggagctg gcacgaatgc tgcgtgagta 1800
ccaggagctc atgagtgtga agctggcctt ggacatcgag atcgccacct accgcaaact 1860
gctggagggc gaggagtaca gaatgtctgg agaatgccag agtgccgtga gcatctctgt 1920
ggtcagcggt agcaccagca ctggaggcat cagcggagga ttaggaagtg gctccgggtt 1980
tggcctgagt agtggctttg gctccggctc tggaagtggc tttgggtttg gtggcagtgt 2040
ctctggcagt tccagcagca agatcatctc taccaccacc ctgaacaaga gacgatagag 2100
gagacgaggt ccctgcagct cactgtgtcc agctgggccc agcactggtg tctctgtgct 2160
tccttcactt cacctccatc ctctgtctct ggggctcatc ttactagtat cccctccact 2220
atcccatggg ctctctctgc cccaggatga tcttctgtgc tgggacaggg actctgcctc 2280
ttggagtttg gtagctactt cttgatttgg gcctggtgac ccacctggaa tgggaaggat 2340
gtcagctgac ctctcacctc ccatgggcag agaagaaaat gaccaggagt gtcatctcca 2400
gaattattgg ggtcacatat gtcccttccc agtccaatgc catctcccac tagatcctgt 2460
attatccatc tacatcagaa ccaaactact tctccaacac ccggcagcac ttggccctgc 2520
aagcttagga tgagaaccac ttagtgtccc attctactcc tctcattccc tcttatccat 2580
ctgcaggtga atcttcaata aaatgctttt gtcattc 2617

<210> 388

<211> 1752

<212> DNA

<213> Homo sapiens

<400> 388
ctgctccttc taggatctcc gcctggttcg gcccgcctgc ctccactcct gcctccacca 60
tgtccatcag ggtgacccag aagtcctaca aggtgtccac ctctggcccc cgggccttca 120
gcagccgctc ctacacgagt gggcccggtt cccgcatcag ctcctcgagc ttctcccgag 180
tgggcagcag caactttcgc ggtggcctgg gcggcggcta tggtggggcc agcggcatgg 240
gaggcatcac cgcagttacg gtcaaccaga gcctgctgag ccccctttgtc ctggaggtgg 300
accccaacat ccaggccgtg cgcacccagg agaaggagca gatcaagacc ctcaacaaca 360
agtttgcctc cttcatagac aaggtacggt tcctggagca gcagaacaag atgctggaga 420
ccaagtggag cctcctgcag cagcagaaga cggctcgaag caacatggac aacatgttcg 480
agagctacat caacaacctt aggcggcagc tggagactct gggccaggag aagctgaagc 540
tggaggcgga gcttggcaac atgcaggggc tggtggagga cttcaagaac aagtatgagg 600
atgagatcaa taagcgtaca gagatggaga acgaatttgt cctcatcaag aaggatgtgg 660
atgaagctta catgaacaag gtagagctgg agtctcgcct ggaagggctg accgacgaga 720
tcaacttcct caggcagcta tatgaagagg agatccggga gctgcagtcc cagatctcgg 780
acacatctgt ggtgctgtcc atggacaaca gccgctccct ggacatggac agcatcattg 840
ctgaggtcaa ggcacagtac gaggatattg ccaaccgcag ccgggctgag gctgagagca 900
tgtaccagat caagtatgag gagctgcaga gcctggctgg gaagcacggg gatgacctgc 960
ggcgcacaaa gactgagatc tctgagatga accggaacat cagccggctc caggctgaga 1020
ttgagggcct caaaggccag agggcttccc tggaggccgc cattgcagat gccgagcagc 1080
gtggagagct ggccattaag gatgccaacg ccaagttgtc cgagctggag gccgccctgc 1140
agcgggccaa gcaggacatg gcgcggcagc tgcgtgagta ccaggagctg atgaacgtca 1200
agctggccct ggacatcgag atcgccacct acaggaagct gctggagggc gaggagagcc 1260
ggctggagtc tgggatgcag aacatgagta ttcatacgaa gaccaccagc ggctatgcag 1320
gtggtctgag ctcggcctat gggggcctca caagccccgg cctcagctac agcctgggct 1380
ccagctttgg ctctggcgcg ggctccagct ccttcagccg caccagctcc tccagggccg 1440
tggttgtgaa gaagatcgag acacgtgatg ggaagctggt gtctgagtcc tctgacgtcc 1500
tgcccaagtg aacagctgcg gcagcccctc ccagcctacc cctcctgcgc tgccccagag 1560
cctgggaagg aggccgctat gcagggtagc actgggaaca ggagacccac ctgaggctca 1620
gccctagccc tcagcccacc tggggagttt actacctggg gacccccctt gcccatgcct 1680
ccagctacaa aacaattcaa ttgctttttt tttttggtcc aaaataaaac ctcagctagc 1740
tctgccaaac cc 1752

<210> 389

<211> 1412

<212> DNA

<213> Homo sapiens

<400> 389
cggggtcgtc cgcaaagcct gagtcctgtc ctttctctct ccccggacag catgagcttc 60
accactcgct ccaccttctc caccaactac cggtccctgg gctctgtcca ggcgcccagc 120
tacggcgccc ggccggtcag cagcgcggcc agcgtctatg caggcgctgg gggctctggt 180
tcccggatct ccgtgtcccg ctccaccagc ttcaggggcg gcatggggtc cgggggcctg 240
gccaccggga tagccggggg tctggcagga atgggaggca tccagaacga gaaggagacc 300
atgcaaagcc tgaacgaccg cctggcctct tacctggaca gagtgaggag cctggagacc 360
gagaaccgga ggctggagag caaaatccgg gagcacttgg agaagaaggg accccaggtc 420
agagactgga gccattactt caagatcatc gaggacctga gggctcagat cttcgcaaat 480
actgtggaca atgcccgcat cgttctgcag attgacaatg cccgtcttgc tgctgatgac 540
tttagagtca agtatgagac agagctggcc atgcgccagt ctgtggagaa cgacatccat 600
gggctccgca aggtcattga tgacaccaat atcacacgac tgcagctgga gacagagatc 660
gaggctctca aggaggagct gctcttcatg aagaagaacc acgaagagga agtaaaaggc 720
ctacaagccc agattgccag ctctgggttg accgtggagg tagatgcccc caaatctcag 780
gacctcgcca agatcatggc agacatccgg gcccaatatg acgagctggc tcggaagaac 840
cgagaggagc tagacaagta ctggtctcag cagattgagg agagcaccac agtggtcacc 900
acacagtctg ctgaggttgg agctgctgag acgacgctca cagagctgag acgtacagtc 960
cagtccttgg agatcgacct ggactccatg agaaatctga aggccagctt ggagaacagc 1020
ctgagggagg tggaggcccg ctacgcccta cagatggagc agctcaacgg gatcctgctg 1080
caccttgagt cagagctggc acagacccgg gcagagggac agcgccaggc ccaggagtat 1140
gaggccctgc tgaacatcaa ggtcaagctg gaggctgaga tcgccaccta ccgccgcctg 1200
ctggaagatg gcgaggactt taatcttggt gatgccttgg acagcagcaa ctccatgcaa 1260
accatccaaa agaccaccac ccgccggata gtggatggca aagtggtgtc tgagaccaat 1320
gacaccaaag ttctgaggca ttaagccagc agaagcaggg taccctttgg ggagcaggag 1380
gccaataaaa agttcagagt tcattggatg tc 1412

<210> 390

<211> 2516

<212> DNA

<213> Homo sapiens

<400> 390
cggggccggc ggacccgcgg gcaggcactg cccgggctgg acgacgtctg gccggctccc 60
ggcgaagggc agcggaggag cggcccagag cgcgcagcta gggcactggc gaaaccccgg 120
gacagtccct ctccgtgcgg gggcggcgca gagcagtccc atccccgggg tcccgggcgc 180
ggctgactgc cggctggttc cctgcgcgca gtagctcccc gagccgggct gcaccggagg 240
cggcgagatg gtcgcgcgcg tcggcctcct gctgcgcgcc ctgcagctgc tactgtgggg 300
ccacctggac gcccagcccg cggagcgcgg aggccaggag ctgcgcaagg aggcggaggc 360
attcctagag aagtacggat acctcaatga acaggtcccc aaagctccca cctccactcg 420
attcagcgat gccatcagag cgtttcagtg ggtgtcccag ctacctgtca gcggcgtgtt 480
ggaccgcgcc accctgcgcc agatgactcg tccccgctgc ggggttacag ataccaacag 540
ttatgcggcc tgggctgaga ggatcagtga cttgtttgct agacaccgga ccaaaatgag 600
gcgtaagaaa cgctttgcaa agcaaggtaa caaatggtac aagcagcacc tctcctaccg 660
cctggtgaac tggcctgagc atctgccgga gccggcagtt cggggcgccg tgcgcgccgc 720
cttccagttg tggagcaacg tctcagcgct ggagttctgg gaggccccag ccacaggccc 780
cgctgacatc cggctcacct tcttccaagg ggaccacaac gatgggctgg gcaatgcctt 840
tgatggccca gggggcgccc tggcgcacgc cttcctgccc cgccgcggcg aagcgcactt 900
cgaccaagat gagcgctggt ccctgagccg ccgccgcggg cgcaacctgt tcgtggtgct 960
ggcgcacgag atcggtcaca cgcttggcct cacccactcg cccgcgccgc gcgcgctcat 1020
ggcgccctac tacaagaggc tgggccgcga cgcgctgctc agctgggacg acgtgctggc 1080
cgtgcagagc ctgtatggga agcccctagg gggctcagtg gccgtccagc tcccaggaaa 1140
gctgttcact gactttgaga cctgggactc ctacagcccc caaggaaggc gccctgaaac 1200
gcagggccct aaatactgcc actcttcctt cgatgccatc actgtagaca ggcaacagca 1260
actgtacatt tttaaaggga gccatttctg ggaggtggca gctgatggca acgtctcaga 1320
gccccgtcca ctgcaggaaa gatgggtcgg gctgcccccc aacattgagg ctgcggcagt 1380
gtcattgaat gatggagatt tctacttctt caaagtgcaa tccgtttgat tgatatgaat 1440
accaaccatc gtgggctcca gcttaaagga gagcagaggg ggtcgatgct ggaggttccg 1500
gggccccaag ccagtgtggg gtctcccaca gctgtgccgg gcaggggccc tgccccgcca 1560
tcctgacgcc gccctcttct tccctcctct gcgccgcctc atcctcttca agggtgcccg 1620

ctactacgtg ctggcccgag ggggactgca agtggagccc tactaccccc gaagtctgca 1680
ggactgggga ggcatccctg aggaggtcag cggcgccctg ccgaggcccg atggctccat 1740
catcttcttc cgagatgacc gctactggcg cctcgaccag gccaaactgc aggcaaccac 1800
ctcgggccgc tgggccaccg agctgccctg gatgggctgc tggcatgcca actcggggag 1860
cgccctgttc tgaaggcacc tcctcacctc agaaactggt ggtgctctca gggcaaaatc 1920
atgttcccca cccccggggc agaacccctc ttagaagcct ctgagtccct ctgcagaaga 1980
ccgggcagca aagcctccat ctggaagtct gtctgccttt gttccttgaa gaatgcagca 2040
ttgtctttgt ctgtccccac cacatggagg tgggggtggg atcaatctta ggaaaagcaa 2100
aaaagggtcc cagatccctt ggccctttcc tccgaggact tctatcctcc ccaggccttt 2160
gtttcttcgg ctaaaggtac agttcctttc aagaggtaac agcactggga tccaagcagg 2220
gggatgaaaa actcagcaga gaaattcgag accattttgc aagactgtgc ccttctcctc 2280
aggacccccct ggctcagttc ttgaaaaacg gtgtcatatt tagtcagagg ccccaccccc 2340
aggaagcatg gatggggatg aaggcacagg cgtctccaac ctcagaggcc ctttgtgggg 2400
tcaggacaca gagtgggagg gagactgatg caggcctacc agtccctggc tttttgtctg 2460
gggctggaat aaagaggtgc cttcagctgg tgggccgaga ggcaggaagc agcctt 2516

<210> 391

<211> 3291

<212> DNA

<213> Homo sapiens

<400> 391
aagtcgcgac cagagccatt ggagggcgcg gggactgcaa ccctaatcag agcccaaatg 60
gcgcagtggg aaatgctgca gaatcttgac agcccctttc aggatcagct gcaccagctt 120
tactcgcaca gcctcctgcc tgtggacatt cgacagtact tggctgtctg gattgaagac 180
cagaactggc aggaagctgc acttgggagt gatgattcca aggctaccat gctattcttc 240
cacttcttgg atcagctgaa ctatgagtgt ggccgttgca gccaggaccc agagtccttg 300
ttgctgcagc acaatttgcg gaaattctgc cgggacattc agccctttc ccaggatcct 360
acccagttgg ctgagatgat ctttaacctc cttctggaag aaaaaagaat tttgatccag 420
gctcagaggg cccaattgga acaaggagag ccagttctcg aaacacctgt ggagagccag 480
caacatgaga ttgaatcccg gatcctggat ttaagggcta tgatggagaa gctggtaaaa 540
tccatcagcc aactgaaaga ccagcaggat gtcttctgct tccgatataa gatccaggcc 600
aaagggaaga caccctctct ggacccccat cagaccaaag agcagaagat tctgcaggaa 660
actctcaatg aactggacaa aaggagaaag gaggtgctgg atgcctccaa agcactgcta 720
ggccgattaa ctaccctaat cgagctactg ctgccaaagt tggaggagtg gaaggcccag 780
cagcaaaaag cctgcatcag agctcccatt gaccacgggt tggaacagct ggagacatgg 840
ttcacagctg gagcaaagct gttgtttcac ctgaggcagc tgctgaagga gctgaaggga 900
ctgagttgcc tggttagcta tcaggatgac cctctgacca aaggggtgga cctacgcaac 960
gcccaggtca cagagttgct acagcgtctg ctccacagag cctttgtggt agaaacccag 1020
ccctgcatgc cccaaactcc ccatcgaccc ctcatcctca agactggcag caagttcacc 1080
gtccgaacaa ggctgctggt gagactccag gaaggcaatg agtcactgac tgtggaagtc 1140
tccattgaca ggaatcctcc tcaattacaa ggcttccgga agttcaacat tctgacttca 1200
aaccagaaaa ctttgacccc cgagaagggg cagagtcagg gtttgatttg ggactttggt 1260
tacctgactc tggtggagca acgttcaggt ggttcaggaa agggcagcaa taaggggcca 1320
ctaggtgtga cagaggaact gcacatcatc agcttcacgg tcaaatatac ctaccagggt 1380
ctgaagcagg agctgaaaac ggacaccctc cctgtggtga ttatttccaa catgaaccag 1440
ctctcaattg cctgggcttc agttctctgg ttcaatttgc tcagcccaaa ccttcagaac 1500
cagcagttct tctccaaccc ccccaaggcc ccctggagct tgctgggccc tgctctcagt 1560
tggcagttct cctcctatgt tggccgaggc ctcaactcag accagctgag catgctgaga 1620
aacaagctgt tcgggcagaa ctgtaggact gaggatccat tattgtcctg ggctgacttc 1680
actaagcgag agagccctcc tggcaagtta ccattctgga catggctgga caaaattctg 1740
gagttggtac atgaccacct gaaggatctc tggaatgatg gacgcatcat gggctttgtg 1800
agtcggagcc aggagcgccg gctgctgaag aagaccatgt ctggcacctt tctactgcgc 1860
ttcagtgaat cgtcagaagg gggcattacc tgctcctggg tggagcacca ggatgatgac 1920
aaggtgctca tctactctgt gcaaccgtac acgaaggagg tgctgcagtc actcccgctg 1980
actgaaatca tccgccatta ccagttgctc actgaggaga atatacctga aaacccactg 2040
cgcttcctct atccccgaat cccccgggat gaagcttttg ggtgctacta ccaggagaaa 2100
gttaatctcc aggaacggag gaaatacctg aaacacaggc tcattgtggt ctctaataga 2160
caggtggatg aactgcaaca accgctggag cttaagccag agccagagct ggagtcatta 2220
gagctggaac tagggctggt gccagagcca gagctcagcc tggacttaga gccactgctg 2280
aaggcagggc tggatctggg gccagagcta gagtctgtgc tggagtccac tctggagcct 2340
gtgatagagc ccacactatg catggtatca caaacagtgc cagagccaga ccaaggacct 2400

```
gtatcacagc cagtgccaga gccagatttg ccctgtgatc tgagacattt gaacactgag   2460
ccaatggaaa tcttcagaaa ctgtgtaaag attgaagaaa tcatgccgaa tggtgaccca   2520
ctgttggctg gccagaacac cgtggatgag gtttacgtct cccgccccag ccacttctac   2580
actgatggac ccttgatgcc ttctgacttc taggaaccac atttcctctg ttcttttcat   2640
atctctttgc ccttcctact cctcatagca tgatattgtt ctccaaggat gggaatcagg   2700
catgtgtccc ttccaagctg tgttaactgt tcaaactcag gcctgtgtga ctccattggg   2760
gtgagaggtg aaagcataac atgggtacag aggggacaac aatgaatcag aacagatgct   2820
gagccatagg tctaaatagg atcctggagg ctgcctgctg tgctgggagg tataggggtc   2880
ctgggggcag gccagggcag ttgacaggta cttggagggc tcagggcagt ggcttctttc   2940
cagtatggaa ggatttcaac attttaatag ttggttaggc taaactggtg catactggca   3000
ttggccttgg tggggagcac agacacagga taggactcca tttctttctt ccattccttc   3060
atgtctagga taacttgctt tcttctttcc tttactcctg gctcaagccc tgaatttctt   3120
cttttcctgc aggggttgag agctttctgc cttagcctac catgtgaaac tctaccctga   3180
agaaagggat ggataggaag tagacctctt tttcttacca gtctcctccc ctactctgcc   3240
ccctaagctg gctgtacctg ttcctccccc ataaaatgat cctgccaatc t            3291
```

<210> 392

<211> 1283

<212> DNA

<213> Homo sapiens

<400> 392

```
ctctctgctc ctcctgttcg acagtcagcc gcatcttctt ttgcgtcgcc agccgagcca     60
catcgctcag acaccatggg gaaggtgaag gtcggagtca acggatttgg tcgtattggg    120
cgcctggtca ccagggctgc ttttaactct ggtaaagtgg atattgttgc catcaatgac    180
cccttcattg acctcaacta catggtttac atgttccaat atgattccac ccatggcaaa    240
ttccatggca ccgtcaaggc tgagaacggg aagcttgtca tcaatggaaa tcccatcacc    300
atcttccagg agcgagatcc ctccaaaatc aagtggggcg atgctggcgc tgagtacgtc    360
gtggagtcca ctggcgtctt caccaccatg gagaaggctg gggctcattt gcagggggga    420
gccaaaaggg tcatcatctc tgccccctct gctgatgccc ccatgttcgt catgggtgtg    480
aaccatgaga agtatgacaa cagcctcaag atcatcagca atgcctcctg caccaccaac    540
tgcttagcac ccctggccaa ggtcatccat gacaactttg gtatcgtgga aggactcatg    600
accacagtcc atgccatcac tgccacccag aagactgtgg atggcccctc cgggaaactg    660
tggcgtgatg gccgcggggc tctccagaac atcatccctg cctctactgg cgctgccaag    720
gctgtgggca aggtcatccc tgagctgaac gggaagctca ctggcatggc cttccgtgtc    780
cccactgcca acgtgtcagt ggtggacctg acctgccgtc tagaaaaacc tgccaaatat    840
gatgacatca agaaggtggt gaagcaggcg tcggagggcc ccctcaaggg catcctgggc    900
tacactgagc accaggtggt ctcctctgac ttcaacagcg acacccactc ctccaccttt    960
gacgctgggg ctggcattgc cctcaacgac cactttgtca agctcatttc ctggtatgac   1020
aacgaatttg gctacagcaa cagggtggtg gacctcatgg cccacatggc ctccaaggag   1080
taagacccct ggaccaccag ccccagcaag agcacaagag gaagagagag accctcactg   1140
ctggggagtc cctgccacac tcagtccccc accacactga atctcccctc ctcacagttg   1200
ccatgtagac cccttgaaga ggggaggggc ctagggagcc gcaccttgtc atgtaccatc   1260
aataaagtac cctgtgctca acc                                           1283
```

<210> 393

<211> 331

<212> PRT

<213> Homo sapiens

<400> 393

```
Met Gly Gly Ser Ala Gly Arg Glu Leu Asp Ala Gly Arg Lys Pro Lys
1               5                   10                  15
Leu Thr Arg Thr Gln Ser Ala Phe Ser Pro Val Ser Phe Ser Pro Leu
            20                  25                  30
Phe Thr Gly Glu Thr Val Ser Leu Val Asp Val Asp Ile Ser Gln Arg
        35                  40                  45
```

Gly Leu Thr Ser Pro His Pro Pro Thr Pro Pro Pro Pro Pro Arg Arg
50 55 60
Ser Leu Ser Leu Leu Asp Asp Ile Ser Gly Thr Leu Pro Thr Ser Val
65 70 75 80
Leu Val Ala Pro Met Gly Ser Ser Leu Gln Ser Phe Pro Leu Pro Pro
85 90 95
Pro Pro Pro Pro His Ala Pro Asp Ala Phe Pro Arg Ile Ala Pro Ile
100 105 110
Arg Ala Ala Glu Ser Leu His Ser Gln Pro Pro Gln His Leu Gln Cys
115 120 125
Pro Leu Tyr Arg Pro Asp Ser Ser Ser Phe Ala Ala Ser Leu Arg Glu
130 135 140
Leu Glu Lys Cys Gly Trp Tyr Trp Gly Pro Met Asn Trp Glu Asp Ala
145 150 155 160
Glu Met Lys Leu Lys Gly Lys Pro Asp Gly Ser Phe Leu Val Arg Asp
165 170 175
Ser Ser Asp Pro Arg Tyr Ile Leu Ser Leu Ser Phe Arg Ser Gln Gly
180 185 190
Ile Thr His His Thr Arg Met Glu His Tyr Arg Gly Thr Phe Ser Leu
195 200 205
Trp Cys His Pro Lys Phe Glu Asp Arg Cys Gln Ser Val Val Glu Phe
210 215 220
Ile Lys Arg Ala Ile Met His Ser Lys Asn Gly Lys Phe Leu Tyr Phe
225 230 235 240
Leu Arg Ser Arg Val Pro Gly Leu Pro Pro Thr Pro Val Gln Leu Leu
245 250 255
Tyr Pro Val Ser Arg Phe Ser Asn Val Lys Ser Leu Gln His Leu Cys
260 265 270
Arg Phe Arg Ile Arg Gln Leu Val Arg Ile Asp His Ile Pro Asp Leu
275 280 285
Pro Leu Pro Lys Pro Leu Ile Ser Tyr Ile Arg Lys Phe Tyr Tyr Tyr
290 295 300
Asp Pro Gln Glu Glu Val Tyr Leu Ser Leu Lys Glu Ala Gln Leu Ile
305 310 315 320
Ser Lys Gln Lys Gln Glu Val Glu Pro Ser Thr
325 330

<210> 394

<211> 306

<212> PRT

<213> Homo sapiens

<400> 394
Met Ala Ala Pro Ile Pro Gln Gly Phe Ser Cys Leu Ser Arg Val Leu
1 5 10 15
Gly Trp Trp Ser Arg Gln Pro Val Leu Val Thr Gln Ser Ala Ala Ile
20 25 30
Val Pro Val Arg Thr Lys Lys Arg Phe Thr Pro Pro Ile Tyr Gln Pro
35 40 45
Lys Phe Lys Thr Glu Lys Glu Phe Met Gln His Ala Arg Lys Ala Gly
50 55 60
Leu Val Ile Pro Pro Glu Lys Ser Asp Arg Ser Ile His Leu Ala Cys
65 70 75 80
Thr Ala Gly Ile Phe Asp Ala Tyr Val Pro Pro Glu Gly Asp Ala Arg
85 90 95
Ile Ser Ser Leu Ser Lys Glu Gly Leu Ile Glu Arg Thr Glu Arg Met
100 105 110
Lys Lys Thr Met Ala Ser Gln Val Ser Ile Arg Arg Ile Lys Asp Tyr
115 120 125
Asp Ala Asn Phe Lys Ile Lys Asp Phe Pro Glu Lys Ala Lys Asp Ile
130 135 140

Phe Ile Glu Ala His Leu Cys Leu Asn Asn Ser Asp His Asp Arg Leu
145 150 155 160
His Thr Leu Val Thr Glu His Cys Phe Pro Asp Met Thr Trp Asp Ile
165 170 175
Lys Tyr Lys Thr Val Arg Trp Ser Phe Val Glu Ser Leu Glu Pro Ser
180 185 190
His Val Val Gln Val Arg Cys Ser Ser Met Met Asn Gln Gly Asn Val
195 200 205
Tyr Gly Gln Ile Thr Val Arg Met His Thr Arg Gln Thr Leu Ala Ile
210 215 220
Tyr Asp Arg Phe Gly Arg Leu Met Tyr Gly Gln Glu Asp Val Pro Lys
225 230 235 240
Asp Val Leu Glu Tyr Val Val Phe Glu Lys Gln Leu Thr Asn Pro Tyr
245 250 255
Gly Ser Trp Arg Met His Thr Lys Ile Val Pro Pro Trp Ala Pro Pro
260 265 270
Lys Gln Pro Ile Leu Lys Thr Val Met Ile Pro Gly Pro Gln Leu Lys
275 280 285
Pro Glu Glu Glu Tyr Glu Glu Ala Gln Gly Glu Ala Gln Lys Pro Gln
290 295 300
Leu Ala
305
<210> 395

<211> 557

<212> PRT

<213> Homo sapiens

<400> 395
Met Val Ser Lys Leu Thr Ser Leu Gln Gln Glu Leu Leu Ser Ala Leu
1 5 10 15
Leu Ser Ser Gly Val Thr Lys Glu Val Leu Val Gln Ala Leu Glu Glu
20 25 30
Leu Leu Pro Ser Pro Asn Phe Gly Val Lys Leu Glu Thr Leu Pro Leu
35 40 45
Ser Pro Gly Ser Gly Ala Glu Pro Asp Thr Lys Pro Val Phe His Thr
50 55 60
Leu Thr Asn Gly His Ala Lys Gly Arg Leu Ser Gly Asp Glu Gly Ser
65 70 75 80
Glu Asp Gly Asp Asp Tyr Asp Thr Pro Pro Ile Leu Lys Glu Leu Gln
85 90 95
Ala Leu Asn Thr Glu Glu Ala Ala Glu Gln Arg Ala Glu Val Asp Arg
100 105 110
Met Leu Ser Glu Asp Pro Trp Arg Ala Ala Lys Met Ile Lys Gly Tyr
115 120 125
Met Gln Gln His Asn Ile Pro Gln Arg Glu Val Val Asp Val Thr Gly
130 135 140
Leu Asn Gln Ser His Leu Ser Gln His Leu Asn Lys Gly Thr Pro Met
145 150 155 160
Lys Thr Gln Lys Arg Ala Ala Leu Tyr Thr Trp Tyr Val Arg Lys Gln
165 170 175
Arg Glu Ile Leu Arg Gln Phe Asn Gln Thr Val Gln Ser Ser Gly Asn
180 185 190
Met Thr Asp Lys Ser Ser Gln Asp Gln Leu Leu Phe Leu Phe Pro Glu
195 200 205
Phe Ser Gln Gln Ser His Gly Pro Gly Gln Ser Asp Asp Ala Cys Ser
210 215 220
Glu Pro Thr Asn Lys Lys Met Arg Arg Asn Arg Phe Lys Trp Gly Pro
225 230 235 240
Ala Ser Gln Gln Ile Leu Tyr Gln Ala Tyr Asp Arg Gln Lys Asn Pro
245 250 255

Ser Lys Glu Glu Arg Glu Ala Leu Val Glu Glu Cys Asn Arg Ala Glu
260 265 270
Cys Leu Gln Arg Gly Val Ser Pro Ser Lys Ala His Gly Leu Gly Ser
275 280 285
Asn Leu Val Thr Glu Val Arg Val Tyr Asn Trp Phe Ala Asn Arg Arg
290 295 300
Lys Glu Glu Ala Phe Arg Gln Lys Leu Ala Met Asp Ala Tyr Ser Ser
305 310 315 320
Asn Gln Thr His Ser Leu Asn Pro Leu Leu Ser His Gly Ser Pro His
325 330 335
His Gln Pro Ser Ser Ser Pro Pro Asn Lys Leu Ser Gly Val Arg Tyr
340 345 350
Ser Gln Gln Gly Asn Asn Glu Ile Thr Ser Ser Ser Thr Ile Ser His
355 360 365
His Gly Asn Ser Ala Met Val Thr Ser Gln Ser Val Leu Gln Gln Val
370 375 380
Ser Pro Ala Ser Leu Asp Pro Gly His Asn Leu Leu Ser Pro Asp Gly
385 390 395 400
Lys Met Ile Ser Val Ser Gly Gly Gly Leu Pro Pro Val Ser Thr Leu
405 410 415
Thr Asn Ile His Ser Leu Ser His His Asn Pro Gln Gln Ser Gln Asn
420 425 430
Leu Ile Met Thr Pro Leu Ser Gly Val Met Ala Ile Ala Gln Ser Leu
435 440 445
Asn Thr Ser Gln Ala Gln Ser Val Pro Val Ile Asn Ser Val Ala Gly
450 455 460
Ser Leu Ala Ala Leu Gln Pro Val Gln Phe Ser Gln Gln Leu His Ser
465 470 475 480
Pro His Gln Gln Pro Leu Met Gln Gln Ser Pro Gly Ser His Met Ala
485 490 495
Gln Gln Pro Phe Met Ala Ala Val Thr Gln Leu Gln Asn Ser His Met
500 505 510
Tyr Ala His Lys Gln Glu Pro Pro Gln Tyr Ser His Thr Ser Arg Phe
515 520 525
Pro Ser Ala Met Val Val Thr Asp Thr Ser Ser Ile Ser Thr Leu Thr
530 535 540
Asn Met Ser Ser Ser Lys Gln Cys Pro Leu Gln Ala Trp
545 550 555
<210> 396

<211> 491

<212> PRT

<213> Homo sapiens

<400> 396
Met Ser Ser Val Glu Ala Lys Ile Glu Asp Lys Lys Val Gln Arg Glu
1 5 10 15
Ser Lys Leu Thr Ser Gly Lys Leu Glu Asn Leu Arg Lys Glu Lys Ile
20 25 30
Asn Phe Leu Arg Asn Lys His Lys Ile His Val Gln Gly Thr Asp Leu
35 40 45
Pro Asp Pro Ile Ala Thr Phe Gln Gln Leu Asp Gln Glu Tyr Lys Ile
50 55 60
Asn Ser Arg Leu Leu Gln Asn Ile Leu Asp Ala Gly Phe Gln Met Pro
65 70 75 80
Thr Pro Ile Gln Met Gln Ala Ile Pro Val Met Leu His Gly Arg Glu
85 90 95
Leu Leu Ala Ser Ala Pro Thr Gly Ser Gly Lys Thr Leu Ala Phe Ser
100 105 110
Ile Pro Ile Leu Met Gln Leu Lys Gln Pro Ala Asn Lys Gly Phe Arg
115 120 125

Ala Leu Ile Ile Ser Pro Thr Arg Glu Leu Ala Ser Gln Ile His Arg
130 135 140
Glu Leu Ile Lys Ile Ser Glu Gly Thr Gly Phe Arg Ile His Met Ile
145 150 155 160
His Lys Ala Ala Val Ala Ala Lys Lys Phe Gly Pro Lys Ser Ser Lys
165 170 175
Lys Phe Asp Ile Leu Val Thr Thr Pro Asn Arg Leu Ile Tyr Leu Leu
180 185 190
Lys Gln Asp Pro Pro Gly Ile Asp Leu Ala Ser Val Glu Trp Leu Val
195 200 205
Val Asp Glu Ser Asp Lys Leu Phe Glu Asp Gly Lys Thr Gly Phe Arg
210 215 220
Asp Gln Leu Ala Ser Ile Phe Leu Ala Cys Thr Ser His Lys Val Arg
225 230 235 240
Arg Ala Met Phe Ser Ala Thr Phe Ala Tyr Asp Val Glu Gln Trp Cys
245 250 255
Lys Leu Asn Leu Asp Asn Val Ile Ser Val Ser Ile Gly Ala Arg Asn
260 265 270
Ser Ala Val Glu Thr Val Glu Gln Glu Leu Leu Phe Val Gly Ser Glu
275 280 285
Thr Gly Lys Leu Leu Ala Val Arg Glu Leu Val Lys Lys Gly Phe Asn
290 295 300
Pro Pro Val Leu Val Phe Val Gln Ser Ile Glu Arg Ala Lys Glu Leu
305 310 315 320
Phe His Glu Leu Ile Tyr Glu Gly Ile Asn Val Asp Val Ile His Ala
325 330 335
Glu Arg Thr Gln Gln Gln Arg Asp Asn Thr Val His Ser Phe Arg Ala
340 345 350
Gly Lys Ile Trp Val Leu Ile Cys Thr Ala Leu Leu Ala Arg Gly Ile
355 360 365
Asp Phe Lys Gly Val Asn Leu Val Ile Asn Tyr Asp Phe Pro Thr Ser
370 375 380
Ser Val Glu Tyr Ile His Arg Ile Gly Arg Thr Gly Arg Ala Gly Asn
385 390 395 400
Lys Gly Lys Ala Ile Thr Phe Phe Thr Glu Asp Asp Lys Pro Leu Leu
405 410 415
Arg Ser Val Ala Asn Val Ile Gln Gln Ala Gly Cys Pro Val Pro Glu
420 425 430
Tyr Ile Lys Gly Phe Gln Lys Leu Leu Ser Lys Gln Lys Lys Lys Met
435 440 445
Ile Lys Lys Pro Leu Glu Arg Glu Ser Ile Ser Thr Thr Pro Lys Cys
450 455 460
Phe Leu Glu Lys Ala Lys Asp Lys Gln Arg Lys Val Thr Gly Gln Asn
465 470 475 480
Ser Lys Lys Lys Val Ala Leu Glu Asp Lys Ser
485 490

<210> 397

<211> 424

<212> PRT

<213> Homo sapiens

<400> 397
Met Asp Phe Ser Arg Arg Ser Phe His Arg Ser Leu Ser Ser Ser Leu
1 5 10 15
Gln Ala Pro Val Val Ser Thr Val Gly Met Gln Arg Leu Gly Thr Thr
20 25 30
Pro Ser Val Tyr Gly Gly Ala Gly Gly Arg Gly Ile Arg Ile Ser Asn
35 40 45
Ser Arg His Thr Val Asn Tyr Gly Ser Asp Leu Thr Gly Gly Gly Asp
50 55 60

Leu Phe Val Gly Asn Glu Lys Met Ala Met Gln Asn Leu Asn Asp Arg
65 70 75 80
Leu Ala Ser Tyr Leu Glu Lys Val Arg Thr Leu Glu Gln Ser Asn Ser
85 90 95
Lys Leu Glu Val Gln Ile Lys Gln Trp Tyr Glu Thr Asn Ala Pro Arg
100 105 110
Ala Gly Arg Asp Tyr Ser Ala Tyr Tyr Arg Gln Ile Glu Glu Leu Arg
115 120 125
Ser Gln Ile Lys Asp Ala Gln Leu Gln Asn Ala Arg Cys Val Leu Gln
130 135 140
Ile Asp Asn Ala Lys Leu Ala Ala Glu Asp Phe Arg Leu Lys Tyr Glu
145 150 155 160
Thr Glu Arg Gly Ile Arg Leu Thr Val Glu Ala Asp Leu Gln Gly Leu
165 170 175
Asn Lys Val Phe Asp Asp Leu Thr Leu His Lys Thr Asp Leu Glu Ile
180 185 190
Gln Ile Glu Glu Leu Asn Lys Asp Leu Ala Leu Leu Lys Lys Glu His
195 200 205
Gln Glu Glu Val Asp Gly Leu His Lys His Leu Gly Asn Thr Val Asn
210 215 220
Val Glu Val Asp Ala Ala Pro Gly Leu Asn Leu Gly Val Ile Met Asn
225 230 235 240
Glu Met Arg Gln Lys Tyr Glu Val Met Ala Gln Lys Asn Leu Gln Glu
245 250 255
Ala Lys Glu Gln Phe Glu Arg Gln Thr Ala Val Leu Gln Gln Gln Val
260 265 270
Thr Val Asn Thr Glu Glu Leu Lys Gly Thr Glu Val Gln Leu Thr Glu
275 280 285
Leu Arg Arg Thr Ser Gln Ser Leu Glu Ile Glu Leu Gln Ser His Leu
290 295 300
Ser Met Lys Glu Ser Leu Glu His Thr Leu Glu Glu Thr Lys Ala Arg
305 310 315 320
Tyr Ser Ser Gln Leu Ala Asn Leu Gln Ser Leu Leu Ser Ser Leu Glu
325 330 335
Ala Gln Leu Met Gln Ile Arg Ser Asn Met Glu Arg Gln Asn Asn Glu
340 345 350
Tyr His Ile Leu Leu Asp Ile Lys Thr Arg Leu Glu Gln Glu Ile Ala
355 360 365
Thr Tyr Arg Arg Leu Leu Glu Gly Glu Asp Val Lys Thr Thr Glu Tyr
370 375 380
Gln Leu Ser Thr Leu Glu Glu Arg Asp Ile Lys Lys Thr Arg Lys Ile
385 390 395 400
Lys Thr Val Val Gln Glu Val Val Asp Gly Lys Val Val Ser Ser Glu
405 410 415
Val Lys Glu Val Glu Glu Asn Ile
420

<210> 398

<211> 209

<212> PRT

<213> Homo sapiens

<400> 398
Met Glu Lys His His Val Pro Ser Asp Phe Asn Val Asn Val Lys Val
1 5 10 15
Asp Thr Gly Pro Arg Glu Asp Leu Ile Lys Val Leu Glu Asp Met Arg
20 25 30
Gln Glu Tyr Glu Leu Ile Ile Lys Lys Lys His Arg Asp Leu Asp Thr
35 40 45
Trp Tyr Lys Glu Gln Ser Ala Ala Met Ser Gln Glu Ala Ala Ser Pro
50 55 60

Ala Thr Val Gln Ser Arg Gln Gly Asp Ile His Glu Leu Lys Arg Thr
65 70 75 80
Phe Gln Ala Leu Glu Ile Asp Leu Gln Thr Gln Tyr Ser Thr Lys Ser
85 90 95
Ala Leu Glu Asn Met Leu Ser Glu Thr Gln Ser Arg Tyr Ser Cys Lys
100 105 110
Leu Gln Asp Met Gln Glu Ile Ile Ser His Tyr Glu Glu Glu Leu Thr
115 120 125
Gln Leu Arg His Glu Leu Glu Arg Gln Asn Asn Glu Tyr Gln Val Leu
130 135 140
Leu Gly Ile Lys Thr His Leu Glu Lys Glu Ile Thr Thr Tyr Arg Arg
145 150 155 160
Leu Leu Glu Gly Glu Ser Glu Gly Thr Arg Glu Glu Ser Lys Ser Ser
165 170 175
Met Lys Val Ser Ala Thr Pro Lys Ile Lys Ala Ile Thr Gln Glu Thr
180 185 190
Ile Asn Gly Arg Leu Val Leu Cys Gln Val Asn Glu Ile Gln Lys His
195 200 205
Ala

<210> 399

<211> 98

<212> PRT

<213> Homo sapiens

<400> 399
Met Asp Cys Cys Ala Ser Arg Gly Cys Ser Val Pro Thr Gly Pro Ala
1 5 10 15
Thr Thr Ile Cys Ser Ser Asp Lys Ser Cys Arg Cys Gly Val Cys Leu
20 25 30
Pro Ser Thr Cys Pro His Thr Val Trp Leu Leu Glu Pro Thr Cys Cys
35 40 45
Asp Asn Cys Pro Pro Pro Cys His Ile Pro Gln Pro Cys Val Pro Thr
50 55 60
Cys Phe Leu Leu Asn Ser Cys Gln Pro Thr Pro Gly Leu Glu Thr Leu
65 70 75 80
Asn Leu Thr Thr Phe Thr Gln Pro Cys Cys Glu Pro Cys Leu Pro Arg
85 90 95
Gly Cys

<210> 400

<211> 98

<212> PRT

<213> Homo sapiens

<400> 400
Met Asp Cys Cys Ala Ser Arg Ser Cys Ser Val Pro Thr Gly Pro Ala
1 5 10 15
Thr Thr Ile Cys Ser Ser Asp Lys Ser Cys Arg Cys Gly Val Cys Leu
20 25 30
Pro Ser Thr Cys Pro His Thr Val Trp Leu Leu Glu Pro Ile Cys Cys
35 40 45
Asp Asn Cys Pro Pro Pro Cys His Ile Pro Gln Pro Cys Val Pro Thr
50 55 60
Cys Phe Leu Leu Asn Ser Cys Gln Pro Thr Pro Gly Leu Glu Thr Leu
65 70 75 80

Asn Leu Thr Thr Phe Thr Gln Pro Cys Cys Glu Pro Cys Leu Pro Arg
85 90 95
Gly Cys

<210> 401

<211> 79

<212> PRT

<213> Homo sapiens

<400> 401
Met Ser Cys Cys Asp Ser Tyr Leu Gln Gly Cys Cys Ser Val Pro Thr
1 5 10 15
Gly Leu Ala Thr Thr Ile Cys Pro Ser Asp Ile Ser Cys Gln Cys Glu
20 25 30
Val Cys Leu Pro Ser Thr Cys Pro His Glu Ile Ser Leu Leu Gln Pro
35 40 45
Thr Cys Cys Glu Pro Gly Pro Cys Leu Ala Ala Cys Leu Thr Pro Met
50 55 60
Cys His Pro Val Asp Cys Ser Thr Asn Ala Thr Gln Leu Gln Pro
65 70 75

<210> 402

<211> 98

<212> PRT

<213> Homo sapiens

<400> 402
Met Tyr Cys Cys Ala Leu Arg Ser Cys Ser Val Pro Thr Gly Pro Ala
1 5 10 15
Thr Thr Phe Cys Ser Phe Asp Lys Ser Cys Arg Cys Gly Val Cys Leu
20 25 30
Pro Ser Thr Cys Pro His Glu Ile Ser Leu Leu Gln Pro Ile Cys Cys
35 40 45
Asp Thr Cys Pro Pro Pro Cys Cys Lys Pro Asp Thr Tyr Val Pro Thr
50 55 60
Cys Trp Leu Leu Asn Asn Cys His Pro Thr Pro Gly Leu Ser Gly Ile
65 70 75 80
Asn Leu Thr Thr Tyr Val Gln Pro Gly Cys Glu Ser Pro Cys Glu Pro
85 90 95
Arg Cys

<210> 403

<211> 174

<212> PRT

<213> Homo sapiens

<400> 403
Met Thr Cys Cys Gln Thr Ser Phe Cys Gly Tyr Pro Ser Phe Ser Ile
1 5 10 15
Ser Gly Thr Cys Gly Ser Ser Cys Cys Gln Pro Ser Cys Cys Glu Thr
20 25 30

Ser Cys Cys Gln Pro Arg Ser Cys Gln Thr Ser Phe Cys Gly Phe Pro
35 40 45
Ser Phe Ser Thr Ser Gly Thr Cys Ser Ser Ser Cys Cys Gln Pro Ser
50 55 60
Cys Cys Glu Thr Ser Cys Cys Gln Pro Ser Cys Cys Glu Thr Ser Cys
65 70 75 80
Cys Gln Pro Ser Cys Cys Gln Ile Ser Ser Cys Gly Thr Gly Cys Gly
85 90 95
Ile Gly Gly Gly Ile Ser Tyr Gly Gln Glu Gly Ser Ser Gly Ala Val
100 105 110
Ser Thr Arg Ile Arg Trp Cys Arg Pro Asp Ser Arg Val Glu Gly Thr
115 120 125
Tyr Leu Pro Pro Cys Cys Val Val Ser Cys Thr Pro Pro Ser Cys Cys
130 135 140
Gln Leu His His Ala Gln Ala Ser Cys Cys Arg Pro Ser Tyr Cys Gly
145 150 155 160
Gln Ser Cys Cys Arg Pro Val Cys Cys Cys Glu Pro Thr Cys
165 170

<210> 404

<211> 167

<212> PRT

<213> Homo sapiens

<400> 404
Met Thr Cys Cys Gln Thr Ser Phe Cys Gly Tyr Pro Ser Cys Ser Thr
1 5 10 15
Ser Gly Thr Cys Gly Ser Ser Cys Cys Gln Pro Ser Cys Cys Glu Thr
20 25 30
Ser Cys Cys Gln Pro Ser Cys Cys Gln Thr Ser Phe Cys Gly Phe Pro
35 40 45
Ser Phe Ser Thr Ser Gly Thr Cys Ser Ser Ser Cys Cys Gln Pro Ser
50 55 60
Cys Cys Glu Thr Ser Cys Cys Gln Pro Ser Cys Cys Gln Thr Ser Ser
65 70 75 80
Cys Gly Thr Gly Cys Gly Ile Gly Gly Gly Ile Gly Tyr Gly Gln Glu
85 90 95
Gly Ser Ser Gly Ala Val Ser Thr Arg Ile Arg Trp Cys Arg Pro Asp
100 105 110
Cys Arg Val Glu Gly Thr Cys Leu Pro Pro Cys Cys Val Val Ser Cys
115 120 125
Thr Pro Pro Thr Cys Cys Gln Leu His His Ala Glu Ala Ser Cys Cys
130 135 140
Arg Pro Ser Tyr Cys Gly Gln Ser Cys Cys Arg Pro Val Cys Cys Cys
145 150 155 160
Tyr Ser Cys Glu Pro Thr Cys
165

<210> 405

<211> 177

<212> PRT

<213> Homo sapiens

<400> 405
Met Ala Cys Cys Gln Thr Ser Phe Cys Gly Phe Pro Ser Cys Ser Thr
1 5 10 15
Ser Gly Thr Cys Gly Ser Ser Cys Cys Gln Pro Ser Cys Cys Glu Thr
20 25 30

Ser Ser Cys Gln Pro Arg Cys Cys Glu Thr Ser Cys Cys Gln Pro Ser
35 40 45
Cys Cys Gln Thr Ser Phe Cys Gly Phe Pro Ser Phe Ser Thr Gly Gly
50 55 60
Thr Cys Asp Ser Ser Cys Cys Gln Pro Ser Cys Cys Glu Thr Ser Cys
65 70 75 80
Cys Gln Pro Ser Cys Tyr Gln Thr Ser Ser Cys Gly Thr Gly Cys Gly
85 90 95
Ile Gly Gly Gly Ile Gly Tyr Gly Gln Glu Gly Ser Ser Gly Ala Val
100 105 110
Ser Thr Arg Ile Arg Trp Cys Arg Pro Asp Cys Arg Val Glu Gly Thr
115 120 125
Cys Leu Pro Pro Cys Cys Val Val Ser Cys Thr Pro Pro Ser Cys Cys
130 135 140
Gln Leu His His Ala Glu Ala Ser Cys Cys Arg Pro Ser Tyr Cys Gly
145 150 155 160
Gln Ser Cys Cys Arg Pro Val Cys Cys Cys Tyr Cys Ser Glu Pro Thr
165 170 175
Cys

<210> 406

<211> 85

<212> PRT

<213> Homo sapiens

<400> 406
Val Thr Cys Val Pro Arg Cys Thr Arg Pro Ile Cys Glu Pro Cys Arg
1 5 10 15
Arg Pro Val Cys Cys Asp Pro Cys Ser Leu Gln Glu Gly Cys Cys Arg
20 25 30
Pro Ile Thr Cys Cys Pro Ser Ser Cys Thr Ala Val Val Cys Arg Pro
35 40 45
Cys Cys Trp Ala Thr Thr Cys Cys Gln Pro Val Ser Val Gln Ser Pro
50 55 60
Cys Cys Arg Pro Pro Cys Gly Gln Pro Thr Pro Cys Ser Thr Thr Cys
65 70 75 80
Arg Thr Ser Ser Cys
85

<210> 407

<211> 128

<212> PRT

<213> Homo sapiens

<400> 407
Met Thr Gly Ser Cys Cys Gly Ser Thr Leu Ser Ser Leu Ser Tyr Gly
1 5 10 15
Gly Gly Cys Cys Gln Pro Cys Cys Cys Arg Asp Pro Cys Cys Arg
20 25 30
Pro Val Thr Cys Gln Thr Thr Val Cys Arg Pro Val Thr Cys Val Pro
35 40 45
Arg Cys Thr Arg Pro Ile Cys Glu Pro Cys Arg Arg Pro Val Cys Cys
50 55 60
Asp Pro Cys Ser Leu Gln Glu Gly Cys Cys Arg Pro Ile Thr Cys Cys
65 70 75 80
Pro Ser Ser Cys Thr Ala Val Val Cys Arg Pro Cys Cys Trp Ala Thr
85 90 95

Thr Cys Cys Gln Pro Val Ser Val Gln Ser Pro Cys Cys Arg Pro Pro
100 105 110
Cys Gly Gln Pro Thr Pro Cys Ser Thr Thr Cys Arg Thr Ser Ser Cys
115 120 125

<210> 408

<211> 20

<212> PRT

<213> Homo sapiens

<400> 408
Met Glu Thr His Cys Thr Gly Arg Ser Ala Ser Phe Cys Ser Ser Ser
1 5 10 15
Ala Ile Leu Ile
20

<210> 409

<211> 210

<212> PRT

<213> Homo sapiens

<400> 409
Met Val Ser Ser Cys Cys Gly Ser Val Cys Ser Asp Gln Gly Cys Gly
1 5 10 15
Gln Val Leu Cys Gln Glu Thr Cys Cys Arg Pro Ser Cys Cys Gln Thr
20 25 30
Thr Cys Cys Arg Thr Thr Cys Tyr Arg Pro Ser Cys Cys Val Ser Ser
35 40 45
Cys Cys Arg Pro Gln Cys Cys Gln Ser Val Cys Cys Gln Pro Thr Cys
50 55 60
Cys Arg Pro Ser Cys Cys Glu Thr Thr Cys Cys His Pro Arg Cys Cys
65 70 75 80
Ile Ser Ser Cys Cys Arg Pro Ser Cys Cys Met Ser Ser Cys Cys Lys
85 90 95
Pro Gln Cys Cys Gln Ser Val Cys Cys Gln Pro Thr Cys Cys Arg Pro
100 105 110
Ser Cys Cys Ile Ser Ser Cys Cys Arg Pro Ser Cys Cys Val Ser Arg
115 120 125
Cys Cys Arg Pro Gln Cys Cys Gln Ser Val Cys Cys Gln Pro Thr Cys
130 135 140
Cys Arg Pro Ser Cys Cys Ile Ser Ser Cys Cys Arg Pro Ser Cys Cys
145 150 155 160
Glu Ser Ser Cys Cys Arg Pro Cys Cys Cys Arg Pro Cys Cys Cys Leu
165 170 175
Arg Pro Val Cys Gly Arg Val Ser Cys His Thr Thr Cys Tyr Arg Pro
180 185 190
Thr Cys Val Ile Ser Thr Cys Pro Arg Pro Leu Cys Cys Ala Ser Ser
195 200 205
Cys Cys
210

<210> 410

<211> 195

<212> PRT

<213> Homo sapiens

<400> 410
Met Val Asn Ser Cys Cys Gly Ser Val Cys Ser His Gln Gly Cys Gly
1 5 10 15
Gln Asp Leu Cys Gln Glu Thr Cys Cys Arg Pro Ser Cys Cys Glu Thr
20 25 30
Thr Cys Cys Arg Thr Thr Tyr Cys Arg Pro Ser Cys Cys Val Ser Ser
35 40 45
Cys Cys Arg Pro Gln Cys Cys Gln Ser Val Cys Cys Gln Pro Thr Cys
50 55 60
Cys Arg Pro Arg Cys Cys Ile Ser Ser Cys Cys Arg Pro Ser Cys Cys
65 70 75 80
Val Ser Ser Cys Cys Lys Pro Gln Cys Cys Gln Ser Met Cys Cys Gln
85 90 95
Pro Thr Cys Cys Arg Pro Arg Cys Cys Ile Ser Ser Cys Cys Arg Pro
100 105 110
Ser Cys Cys Val Ser Ser Cys Cys Arg Pro Gln Cys Cys Gln Ser Val
115 120 125
Cys Cys Gln Pro Thr Cys Cys His Pro Ser Cys Ser Ile Ser Ser Cys
130 135 140
Cys Arg Pro Ser Cys Cys Glu Ser Ser Cys Cys Arg Pro Cys Cys Cys
145 150 155 160
Leu Arg Pro Val Cys Gly Gly Val Ser Cys His Thr Thr Cys Tyr Arg
165 170 175
Pro Thr Cys Val Ile Ser Ser Cys Pro Arg Pro Leu Cys Cys Ala Ser
180 185 190
Ser Cys Cys
195

<210> 411

<211> 201

<212> PRT

<213> Homo sapiens

<400> 411
Met Val Asn Ser Cys Cys Gly Ser Val Cys Ser Asp Gln Gly Cys Gly
1 5 10 15
Leu Glu Asn Cys Cys Arg Pro Ser Cys Cys Gln Thr Thr Cys Cys Arg
20 25 30
Thr Thr Cys Cys Arg Pro Ser Cys Cys Val Ser Ser Cys Cys Arg Pro
35 40 45
Gln Cys Cys Gln Ser Val Cys Cys Gln Pro Thr Cys Cys Arg Pro Ser
50 55 60
Cys Cys Gln Thr Thr Cys Cys Arg Thr Thr Cys Cys Arg Pro Ser Cys
65 70 75 80
Cys Val Ser Ser Cys Cys Arg Pro Gln Cys Cys Gln Ser Val Cys Cys
85 90 95
Gln Pro Thr Cys Cys Arg Pro Ser Cys Cys Gln Thr Thr Cys Cys Arg
100 105 110
Thr Thr Cys Cys Arg Pro Ser Cys Cys Val Ser Ser Cys Cys Arg Pro
115 120 125
Gln Cys Cys Gln Ser Val Cys Cys Gln Pro Thr Cys Cys Arg Pro Ser
130 135 140
Cys Cys Ile Ser Ser Ser Cys Cys Pro Ser Cys Cys Glu Ser Ser Cys
145 150 155 160
Cys Arg Pro Cys Cys Cys Leu Arg Pro Val Cys Gly Arg Val Ser Cys
165 170 175
His Thr Thr Cys Tyr Arg Pro Thr Cys Val Ile Ser Thr Cys Pro Arg
180 185 190
Pro Leu Cys Cys Ala Ser Ser Cys Cys
195 200

<210> 412

<211> 186

<212> PRT

<213> Homo sapiens

<400> 412

Met Val Ser Ser Cys Cys Gly Ser Val Ser Ser Glu Gln Ser Cys Gly
1 5 10 15
Leu Glu Asn Cys Cys Arg Pro Ser Cys Cys Gln Thr Thr Cys Cys Arg
20 25 30
Thr Thr Cys Cys Arg Pro Ser Cys Cys Lys Pro Gln Cys Cys Gln Ser
35 40 45
Val Cys Tyr Gln Pro Thr Cys Cys His Pro Ser Cys Cys Ile Ser Ser
50 55 60
Cys Cys His Pro Tyr Cys Cys Glu Ser Ser Cys Cys Arg Pro Cys Cys
65 70 75 80
Cys Arg Pro Ser Cys Cys Gln Thr Thr Cys Cys Arg Thr Thr Cys Cys
85 90 95
Arg Thr Thr Cys Cys Cys Pro Ser Cys Cys Val Ser Ser Cys Cys Arg
100 105 110
Pro Gln Cys Cys Gln Ser Val Cys Cys Gln Pro Thr Cys Cys Arg Pro
115 120 125
Ser Cys Cys Ile Ser Ser Cys Cys His Pro Ser Cys Cys Glu Ser Ser
130 135 140
Cys Cys Arg Pro Cys Cys Cys Val Arg Pro Val Cys Gly Arg Val Ser
145 150 155 160
Cys His Thr Thr Cys Tyr Arg Pro Thr Cys Val Ile Ser Thr Cys Pro
165 170 175
Arg Pro Leu Cys Cys Ala Ser Ser Cys Cys
180 185

<210> 413

<211> 106

<212> PRT

<213> Homo sapiens

<400> 413

Met Val Asn Ser Cys Cys Gly Ser Val Cys Ser Asp Gln Gly Cys Gly
1 5 10 15
Leu Glu Asn Cys Cys Arg Pro Ser Tyr Cys Gln Thr Thr Cys Cys Arg
20 25 30
Thr Thr Cys Cys Arg Pro Ser Cys Cys Arg Pro Ser Cys Cys Arg Pro
35 40 45
Gln Cys Cys Gln Ser Val Cys Cys Gln Pro Thr Cys Cys Cys Pro Ser
50 55 60
Tyr Cys Val Ser Ser Cys Cys Arg Pro Gln Cys Cys Gln Thr Thr Arg
65 70 75 80
Cys Arg Thr Thr Cys Cys Arg Pro Ser Cys Cys Val Ser Arg Cys Tyr
85 90 95
Arg Pro His Cys Gly Gln Ser Leu Cys Cys
100 105

<210> 414

<211> 166

<212> PRT

<213> Homo sapiens

<400> 414

```
Met Val Asn Ser Cys Cys Gly Ser Val Cys Ser Asp Gln Gly Cys Gly
1               5                   10                  15
Leu Glu Asn Cys Cys Arg Pro Ser Tyr Cys Gln Thr Thr Cys Cys Arg
            20                  25                  30
Thr Thr Cys Cys Arg Pro Ser Cys Cys Val Ser Ser Cys Cys Arg Pro
        35                  40                  45
Gln Cys Cys Gln Thr Thr Cys Cys Arg Thr Thr Cys Cys His Pro Ser
    50                  55                  60
Cys Cys Val Ser Ser Cys Cys Arg Pro Gln Cys Cys Gln Ser Val Cys
65                  70                  75                  80
Cys Gln Pro Thr Cys Cys Arg Pro Gln Cys Cys Gln Thr Thr Cys Cys
                85                  90                  95
Arg Thr Thr Cys Cys Arg Pro Ser Cys Cys Arg Pro Gln Cys Cys Gln
            100                 105                 110
Ser Val Cys Cys Gln Pro Thr Cys Cys Cys Pro Ser Tyr Cys Val Ser
        115                 120                 125
Ser Cys Cys Arg Pro Gln Cys Cys Gln Thr Thr Cys Cys Arg Thr Thr
    130                 135                 140
Cys Cys Arg Pro Ser Cys Cys Val Ser Arg Cys Tyr Arg Pro His Cys
145                 150                 155                 160
Gly Gln Ser Leu Cys Cys
                165
```

<210> 415

<211> 136

<212> PRT

<213> Homo sapiens

<400> 415

```
Met Val Asn Ser Cys Cys Gly Ser Val Cys Ser Asp Gln Gly Cys Gly
1               5                   10                  15
Leu Glu Asn Cys Cys Arg Pro Ser Cys Cys Gln Thr Thr Cys Cys Arg
            20                  25                  30
Thr Thr Cys Cys Arg Pro Ser Cys Cys Val Ser Ser Cys Cys Arg Pro
        35                  40                  45
Gln Cys Cys Gln Ser Val Cys Cys Gln Pro Thr Cys Cys Ser Pro Ser
    50                  55                  60
Cys Cys Gln Thr Thr Cys Cys Arg Thr Thr Cys Cys Arg Pro Ser Cys
65                  70                  75                  80
Cys Val Ser Ser Cys Phe Arg Pro Gln Cys Cys Gln Ser Val Cys Cys
                85                  90                  95
Gln Pro Thr Cys Cys Arg Pro Ser Cys Gly Gln Thr Thr Cys Cys Arg
            100                 105                 110
Thr Thr Cys Tyr Arg Pro Ser Cys Cys Val Ser Thr Cys Cys Arg Pro
        115                 120                 125
Thr Cys Ser Ser Gly Ser Cys Cys
    130                 135
```

<210> 416

<211> 127

<212> PRT

<213> Homo sapiens

<400> 416
Met Val Asn Ser Cys Cys Gly Ser Val Cys Ser Asp Gln Gly Cys Asp
1 5 10 15
Gln Gly Leu Cys Gln Glu Thr Cys Cys Arg Pro Ser Cys Cys Gln Thr
20 25 30
Thr Cys Cys Cys Pro Ser Cys Val Val Ser Ser Cys Cys Arg Pro Ser
35 40 45
Cys Ser Gln Thr Thr Cys Cys Gln Thr Thr Cys Cys Arg Pro Ser Cys
50 55 60
Cys Arg Pro Val Cys Cys Gln Thr Thr Cys Arg Pro Ser Cys Gly Val
65 70 75 80
Ser Ser Cys Cys Arg Pro Leu Cys Cys Gln Thr Thr Cys Arg Pro Ser
85 90 95
Cys Gly Val Ser Ser Cys Cys Arg Pro Leu Cys Cys Gln Thr Thr Cys
100 105 110
Cys Arg Thr Thr Cys Cys Arg Pro Ser Cys Cys Gly Ser Ser Cys
115 120 125

<210> 417

<211> 174

<212> PRT

<213> Homo sapiens

<400> 417
Met Thr His Cys Cys Ser Pro Cys Cys Gln Pro Thr Cys Cys Arg Thr
1 5 10 15
Thr Cys Cys Arg Thr Thr Cys Trp Lys Pro Thr Thr Val Thr Thr Cys
20 25 30
Ser Ser Thr Pro Cys Cys Gln Pro Ala Cys Cys Val Ser Ser Cys Cys
35 40 45
Gln Pro Cys Cys Arg Pro Thr Cys Cys Gln Asn Thr Cys Cys Arg Thr
50 55 60
Thr Cys Cys Gln Pro Thr Cys Val Thr Ser Cys Cys Gln Pro Ser Cys
65 70 75 80
Cys Ser Thr Pro Cys Cys Gln Pro Thr Cys Cys Gly Ser Ser Cys Cys
85 90 95
Gly Gln Thr Ser Cys Gly Ser Ser Cys Gly Gln Ser Ser Ser Cys Ala
100 105 110
Pro Val Tyr Cys Arg Arg Thr Cys Tyr Tyr Pro Thr Thr Val Cys Leu
115 120 125
Pro Gly Cys Leu Asn Gln Ser Cys Gly Ser Asn Cys Cys Gln Pro Cys
130 135 140
Cys Arg Pro Ala Cys Cys Glu Thr Thr Cys Cys Arg Thr Thr Cys Phe
145 150 155 160
Gln Pro Thr Cys Val Ser Ser Cys Cys Gln Pro Ser Cys Cys
165 170

<210> 418

<211> 159

<212> PRT

<213> Homo sapiens

<400> 418
Met Thr His Cys Cys Ser Pro Cys Cys Gln Pro Thr Cys Cys Arg Thr
1 5 10 15
Thr Cys Trp Gln Pro Thr Thr Val Thr Thr Cys Ser Ser Thr Pro Cys
20 25 30

Cys Gln Pro Ser Cys Cys Val Ser Ser Cys Cys Gln Pro Cys Cys His
35 40 45
Pro Thr Cys Cys Gln Asn Thr Cys Cys Arg Thr Thr Cys Cys Gln Pro
50 55 60
Ile Cys Val Thr Ser Cys Cys Gln Pro Ser Cys Cys Ser Thr Pro Cys
65 70 75 80
Cys Gln Pro Thr Cys Cys Gly Ser Ser Cys Gly Gln Ser Ser Ser Cys
85 90 95
Ala Pro Val Tyr Cys Arg Arg Thr Cys Tyr His Pro Thr Ser Val Cys
100 105 110
Leu Pro Gly Cys Leu Asn Gln Ser Cys Gly Ser Asn Cys Cys Gln Pro
115 120 125
Cys Cys Arg Pro Ala Cys Cys Glu Thr Thr Cys Cys Arg Thr Thr Cys
130 135 140
Phe Gln Pro Thr Cys Val Tyr Ser Cys Cys Gln Pro Ser Cys Cys
145 150 155

<210> 419

<211> 159

<212> PRT

<213> Homo sapiens

<400> 419
Met Thr His Cys Cys Ser Pro Cys Cys Gln Pro Thr Cys Cys Arg Thr
1 5 10 15
Thr Cys Trp Lys Pro Thr Thr Val Thr Thr Cys Ser Ser Thr Pro Cys
20 25 30
Cys Gln Pro Ser Cys Cys Val Ser Ser Cys Cys Gln Pro Cys Cys Arg
35 40 45
Pro Thr Cys Cys Gln Asn Thr Cys Cys Gln Pro Ile Cys Val Thr Ser
50 55 60
Cys Cys Gln Pro Ser Cys Cys Ser Thr Pro Cys Cys Gln Pro Thr Cys
65 70 75 80
Cys Gly Gln Thr Ser Cys Gly Ser Ser Cys Gly Gln Ser Ser Ser Cys
85 90 95
Ala Pro Val Tyr Cys Arg Arg Thr Cys Tyr His Pro Thr Thr Val Cys
100 105 110
Leu Pro Gly Cys Leu Asn Gln Ser Cys Gly Ser Ser Cys Cys Gln Pro
115 120 125
Cys Cys Arg Pro Ala Cys Cys Glu Thr Thr Cys Cys Arg Thr Thr Cys
130 135 140
Phe Gln Pro Thr Cys Val Tyr Ser Cys Cys Gln Pro Ser Cys Cys
145 150 155

<210> 420

<211> 154

<212> PRT

<213> Homo sapiens

<400> 420
Met Thr His Cys Cys Ser Pro Cys Cys Gln Pro Thr Cys Cys Arg Thr
1 5 10 15
Thr Cys Cys Arg Thr Thr Cys Trp Lys Pro Thr Thr Val Thr Thr Cys
20 25 30
Ser Ser Thr Pro Cys Cys Gln Pro Ser Cys Cys Val Ser Ser Cys Cys
35 40 45
Gln Pro Cys Cys Arg Pro Ala Cys Cys Gln Asn Thr Cys Cys Arg Thr
50 55 60

Thr Cys Cys Gln Pro Thr Cys Leu Ser Ser Cys Cys Gly Gln Thr Ser
65 70 75 80
Cys Gly Ser Ser Cys Gly Gln Ser Ser Ser Cys Ala Pro Val Tyr Cys
85 90 95
Arg Arg Thr Cys Tyr Tyr Pro Thr Thr Val Cys Leu Pro Gly Cys Leu
100 105 110
Asn Gln Ser Cys Gly Ser Ser Cys Cys Gln Pro Cys Cys Arg Pro Ala
115 120 125
Cys Cys Glu Thr Thr Cys Cys Arg Thr Thr Cys Phe Gln Pro Thr Cys
130 135 140
Val Ser Ser Cys Cys Gln Pro Ser Cys Cys
145 150
<210> 421

<211> 154

<212> PRT

<213> Homo sapiens

<400> 421
Met Thr His Cys Cys Ser Pro Cys Cys Gln Pro Thr Cys Cys Arg Thr
1 5 10 15
Thr Cys Cys Arg Thr Thr Cys Trp Lys Pro Thr Thr Val Thr Thr Cys
20 25 30
Ser Ser Thr Pro Cys Cys Gln Pro Ser Cys Cys Val Ser Ser Cys Cys
35 40 45
Gln Pro Cys Cys Arg Pro Thr Cys Cys Gln Asn Thr Cys Cys Gln Pro
50 55 60
Thr Cys Val Thr Ser Cys Cys Gln Pro Ser Cys Cys Ser Thr Pro Cys
65 70 75 80
Cys Gln Pro Thr Cys Cys Gly Ser Ser Cys Asp Gln Ser Ser Ser Cys
85 90 95
Ala Pro Val Tyr Cys Arg Arg Thr Cys Tyr Tyr Pro Thr Thr Val Cys
100 105 110
Leu Pro Gly Cys Leu Asn Gln Ser Cys Gly Ser Asn Cys Cys Gln Pro
115 120 125
Cys Cys Arg Pro Ala Cys Cys Glu Thr Thr Cys Phe Gln Pro Thr Cys
130 135 140
Val Ser Ser Cys Cys Gln Pro Phe Cys Cys
145 150
<210> 422

<211> 138

<212> PRT

<213> Homo sapiens

<400> 422
Met Leu Gln Asp His Leu Leu Gln Asp Asn Leu Leu Glu Ala His His
1 5 10 15
Cys Asp His Leu Gln Gln His Ile Leu Leu Pro Ala Leu Leu Leu Cys
20 25 30
Val Gln Leu Leu Pro Ala Leu Leu Pro Pro Asn Leu Leu Ser Lys His
35 40 45
Leu Leu Gln Asp His Leu Leu Pro Ala His Leu Cys Asp Gln Leu Leu
50 55 60
Pro Ala Phe Leu Leu Gln His Thr Leu Leu Thr Ala His Leu Leu Trp
65 70 75 80
Val Gln Leu Leu Trp Pro Asn His Leu Trp Val Gln Leu Leu Pro Ala
85 90 95

Gln Leu Leu Cys Thr His Leu Leu Gln Glu Asn Leu Leu Pro Pro His
            100   105   110
Glu Cys Leu Pro Ala Trp Leu Pro Lys Ser Glu Leu Trp Leu Gln Leu
        115   120   125
Leu Pro Ala Leu Leu Pro Pro Ser Leu Leu
    130   135

<210> 423

<211> 409

<212> PRT

<213> Homo sapiens

<400> 423

Met Ser Gly Ser Cys Ser Ser Arg Lys Cys Phe Ser Val Pro Ala Thr
1   5   10   15
Ser Leu Cys Ser Thr Glu Val Ser Cys Gly Gly Pro Ile Cys Leu Pro
            20   25   30
Ser Ser Cys Gln Ser Gln Thr Trp Gln Leu Val Thr Cys Gln Asp Ser
        35   40   45
Cys Gly Ser Ser Ser Cys Gly Pro Gln Cys Arg Gln Pro Ser Cys Pro
    50   55   60
Val Ser Ser Cys Ala Gln Pro Leu Cys Cys Asp Pro Val Ile Cys Glu
65   70   75   80
Pro Ser Cys Ser Val Ser Ser Gly Cys Gln Pro Val Cys Cys Glu Ala
                85   90   95
Thr Thr Cys Glu Pro Ser Cys Ser Val Ser Asn Cys Tyr Gln Pro Val
            100   105   110
Cys Phe Glu Ala Thr Ile Cys Glu Pro Ser Cys Ser Val Ser Asn Cys
        115   120   125
Cys Gln Pro Val Cys Phe Glu Ala Thr Val Cys Glu Pro Ser Cys Ser
    130   135   140
Val Ser Ser Cys Ala Gln Pro Val Cys Cys Glu Pro Ala Ile Cys Glu
145   150   155   160
Pro Ser Cys Ser Val Ser Ser Cys Cys Gln Pro Val Gly Ser Glu Ala
                165   170   175
Thr Ser Cys Gln Pro Val Leu Cys Val Pro Thr Ser Cys Gln Pro Val
            180   185   190
Leu Cys Lys Ser Ser Cys Cys Gln Pro Val Val Cys Glu Pro Ser Cys
        195   200   205
Cys Ser Ala Val Cys Thr Leu Pro Ser Ser Cys Gln Pro Val Val Cys
    210   215   220
Glu Pro Ser Cys Cys Gln Pro Val Cys Pro Thr Pro Thr Cys Ser Val
225   230   235   240
Thr Ser Ser Cys Gln Ala Val Cys Cys Asp Pro Ser Pro Trp Ser Ser
                245   250   255
Ala Ser Ala Ile Cys Arg Pro Thr Cys Pro Arg Thr Phe Tyr Ile Pro
            260   265   270
Ser Ser Ser Lys Arg Pro Cys Ser Ala Thr Ile Ser Tyr Arg Pro Val
        275   280   285
Ser Arg Pro Ile Cys Arg Pro Ile Cys Ser Gly Leu Leu Thr Tyr Arg
    290   295   300
Gln Pro Tyr Met Thr Ser Ile Ser Tyr Arg Pro Ala Cys Tyr Arg Pro
305   310   315   320
Cys Tyr Ser Ile Leu Arg Arg Pro Ala Cys Val Thr Ser Tyr Ser Cys
                325   330   335
Arg Pro Val Tyr Phe Arg Pro Ser Cys Thr Glu Ser Asp Ser Cys Lys
            340   345   350
Arg Asp Cys Lys Lys Ser Thr Ser Ser Gln Leu Asp Cys Val Asp Thr
        355   360   365
Thr Pro Cys Lys Val Asp Val Ser Glu Glu Ala Pro Cys Gln Pro Thr
    370   375   380

Glu Ala Lys Pro Ile Ser Pro Thr Thr Arg Glu Ala Ala Ala Ala Gln
385 390 395 400
Pro Ala Ala Ser Lys Pro Ala Asn Cys
405

<210> 424

<211> 105

<212> PRT

<213> Homo sapiens

<400> 424
Met Gly Cys Cys Pro Gly Asp Cys Phe Thr Cys Cys Thr Gln Glu Gln
1 5 10 15
Asn Cys Cys Glu Glu Cys Cys Cys Gln Pro Gly Cys Cys Gly Cys Cys
20 25 30
Gly Ser Cys Cys Gly Cys Gly Gly Ser Gly Cys Gly Gly Ser Gly Cys
35 40 45
Gly Gly Ser Cys Cys Gly Ser Ser Cys Cys Gly Ser Gly Cys Gly Gly
50 55 60
Cys Gly Gly Cys Gly Gly Cys Gly Gly Gly Cys Cys Gly Ser Ser Cys
65 70 75 80
Cys Gly Ser Ser Cys Cys Gly Ser Gly Cys Cys Gly Pro Val Cys Cys
85 90 95
Gln Pro Thr Pro Ile Cys Asp Thr Lys
100 105

<210> 425

<211> 404

<212> PRT

<213> Homo sapiens

<400> 425
Met Ser Tyr Ser Cys Gly Leu Pro Ser Leu Ser Cys Arg Thr Ser Cys
1 5 10 15
Ser Ser Arg Pro Cys Val Pro Pro Ser Cys His Gly Cys Thr Leu Pro
20 25 30
Gly Ala Cys Asn Ile Pro Ala Asn Val Ser Asn Cys Asn Trp Phe Cys
35 40 45
Glu Gly Ser Phe Asn Gly Ser Glu Lys Glu Thr Met Gln Phe Leu Asn
50 55 60
Asp Arg Leu Ala Ser Tyr Leu Glu Lys Val Arg Gln Leu Glu Arg Asp
65 70 75 80
Asn Ala Glu Leu Glu Asn Leu Ile Arg Glu Arg Ser Gln Gln Gln Glu
85 90 95
Pro Leu Val Cys Ala Ser Tyr Gln Ser Tyr Phe Lys Thr Ile Glu Glu
100 105 110
Leu Gln Gln Lys Ile Leu Cys Ser Lys Ser Glu Asn Ala Arg Leu Val
115 120 125
Val Gln Ile Asp Asn Ala Lys Leu Ala Ser Asp Asp Phe Arg Thr Lys
130 135 140
Tyr Glu Thr Glu Leu Ser Leu Arg Gln Leu Val Glu Ser Asp Ile Asn
145 150 155 160
Gly Leu Arg Arg Ile Leu Asp Glu Leu Thr Leu Cys Arg Ser Asp Leu
165 170 175
Glu Ala Gln Val Glu Ser Leu Lys Glu Glu Leu Leu Cys Leu Lys Gln
180 185 190
Asn His Glu Gln Glu Val Asn Thr Leu Arg Cys Gln Leu Gly Asp Arg
195 200 205

```
Leu Asn Val Glu Val Asp Ala Ala Pro Thr Val Asp Leu Asn Gln Val
    210                 215                 220
Leu Asn Glu Thr Arg Ser Gln Tyr Glu Ala Leu Val Glu Thr Asn Arg
225                 230                 235                 240
Arg Glu Val Glu Gln Trp Phe Ala Thr Gln Thr Glu Glu Leu Asn Lys
                245                 250                 255
Gln Val Val Ser Ser Ser Glu Gln Leu Gln Ser Tyr Gln Ala Glu Ile
            260                 265                 270
Ile Glu Leu Arg Arg Thr Val Asn Ala Leu Glu Ile Glu Leu Gln Ala
        275                 280                 285
Gln His Asn Leu Arg Asp Ser Leu Glu Asn Thr Leu Thr Glu Ser Glu
    290                 295                 300
Ala Arg Tyr Ser Ser Gln Leu Ser Gln Val Gln Arg Leu Ile Thr Asn
305                 310                 315                 320
Val Glu Ser Gln Leu Ala Glu Ile Arg Ser Asp Leu Glu Arg Gln Asn
                325                 330                 335
Gln Glu Tyr Gln Val Leu Leu Asp Val Arg Ala Arg Leu Glu Cys Glu
            340                 345                 350
Ile Asn Thr Tyr Arg Ser Leu Leu Glu Ser Glu Asp Cys Lys Leu Pro
        355                 360                 365
Ser Asn Pro Cys Ala Thr Thr Asn Ala Cys Asp Lys Ser Thr Gly Pro
    370                 375                 380
Cys Ile Ser Asn Pro Cys Gly Leu Arg Ala Arg Cys Gly Pro Cys Asn
385                 390                 395                 400
Thr Phe Gly Tyr
```

<210> 426

<211> 404

<212> PRT

<213> Homo sapiens

<400> 426

```
Met Pro Tyr Asn Phe Cys Leu Pro Ser Leu Ser Cys Arg Thr Ser Cys
1               5                   10                  15
Ser Ser Arg Pro Cys Val Pro Pro Ser Cys His Gly Tyr Thr Leu Pro
            20                  25                  30
Gly Ala Cys Asn Ile Pro Ala Asn Val Ser Asn Cys Asn Trp Phe Cys
        35                  40                  45
Glu Gly Ser Phe Asn Gly Ser Glu Lys Glu Thr Met Gln Phe Leu Asn
    50                  55                  60
Asp Arg Leu Ala Ser Tyr Leu Glu Lys Val Arg Gln Leu Glu Arg Asp
65                  70                  75                  80
Asn Ala Glu Leu Glu Asn Leu Ile Arg Glu Arg Ser Gln Gln Gln Glu
                85                  90                  95
Pro Leu Leu Cys Pro Ser Tyr Gln Ser Tyr Phe Lys Thr Ile Glu Glu
            100                 105                 110
Leu Gln Gln Lys Ile Leu Cys Ser Lys Ser Glu Asn Ala Arg Leu Val
        115                 120                 125
Val Gln Ile Asp Asn Ala Lys Leu Ala Ala Asp Asp Phe Arg Thr Lys
    130                 135                 140
Tyr Gln Thr Glu Gln Ser Leu Arg Gln Leu Val Glu Ser Asp Ile Asn
145                 150                 155                 160
Ser Leu Arg Arg Ile Leu Asp Glu Leu Thr Leu Cys Arg Ser Asp Leu
                165                 170                 175
Glu Ala Gln Met Glu Ser Leu Lys Glu Glu Leu Leu Ser Leu Lys Gln
            180                 185                 190
Asn His Glu Gln Glu Val Asn Thr Leu Arg Cys Gln Leu Gly Asp Arg
        195                 200                 205
Leu Asn Val Glu Val Asp Ala Ala Pro Ala Val Asp Leu Asn Gln Val
    210                 215                 220
```

Leu Asn Glu Thr Arg Asn Gln Tyr Glu Ala Leu Val Glu Thr Asn Arg
225 230 235 240
Arg Glu Val Glu Gln Trp Phe Ala Thr Gln Thr Glu Glu Leu Asn Lys
245 250 255
Gln Val Val Ser Ser Ser Glu Gln Leu Gln Ser Tyr Gln Ala Glu Ile
260 265 270
Ile Glu Leu Arg Arg Thr Val Asn Ala Leu Glu Ile Glu Leu Gln Ala
275 280 285
Gln His Asn Leu Arg Tyr Ser Leu Glu Asn Thr Leu Thr Glu Ser Glu
290 295 300
Ala Arg Tyr Ser Ser Gln Leu Ser Gln Val Gln Ser Leu Ile Thr Asn
305 310 315 320
Val Glu Ser Gln Leu Ala Glu Ile Arg Ser Asp Leu Glu Arg Gln Asn
325 330 335
Gln Glu Tyr Gln Val Leu Leu Asp Val Arg Ala Arg Leu Glu Cys Glu
340 345 350
Ile Asn Thr Tyr Arg Ser Leu Leu Glu Ser Glu Asp Cys Lys Leu Pro
355 360 365
Ser Asn Pro Cys Ala Thr Thr Asn Ala Cys Glu Lys Pro Ile Gly Ser
370 375 380
Cys Val Thr Asn Pro Cys Gly Pro Arg Ser Arg Cys Gly Pro Cys Asn
385 390 395 400
Thr Phe Gly Tyr

<210> 427

<211> 436

<212> PRT

<213> Homo sapiens

<400> 427
Met Leu Tyr Ala Lys Pro Pro Pro Thr Ile Asn Gly Ile Lys Gly Leu
1 5 10 15
Gln Arg Lys Glu Arg Leu Lys Pro Ala His Ile His Leu Gln Gln Leu
20 25 30
Thr Cys Phe Ser Ile Thr Cys Ser Ser Thr Met Ser Tyr Ser Cys Cys
35 40 45
Leu Pro Ser Leu Gly Cys Arg Thr Ser Cys Ser Ser Arg Pro Cys Val
50 55 60
Pro Pro Ser Cys His Gly Tyr Thr Leu Pro Gly Ala Cys Asn Ile Pro
65 70 75 80
Ala Asn Val Ser Asn Cys Asn Trp Phe Cys Glu Gly Ser Phe Asn Gly
85 90 95
Ser Glu Lys Glu Thr Met Gln Phe Leu Asn Asp Arg Leu Ala Ser Tyr
100 105 110
Leu Glu Lys Val Arg Gln Leu Glu Arg Asp Asn Ala Glu Leu Glu Lys
115 120 125
Leu Ile Gln Glu Arg Ser Gln Gln Gln Glu Pro Leu Leu Cys Pro Ser
130 135 140
Tyr Gln Ser Tyr Phe Lys Thr Ile Glu Glu Leu Gln Gln Lys Ile Leu
145 150 155 160
Cys Ala Lys Ala Glu Asn Ala Arg Leu Val Val Asn Ile Asp Asn Ala
165 170 175
Lys Leu Ala Ser Asp Asp Phe Arg Ser Lys Tyr Gln Thr Glu Gln Ser
180 185 190
Leu Arg Leu Leu Val Glu Ser Asp Ile Asn Ser Ile Arg Arg Ile Leu
195 200 205
Asp Glu Leu Thr Leu Cys Lys Ser Asp Leu Glu Ser Gln Val Glu Ser
210 215 220
Leu Arg Glu Glu Leu Ile Cys Leu Lys Lys Asn His Glu Glu Glu Val
225 230 235 240

Asn Thr Leu Arg Ser Gln Leu Gly Asp Arg Leu Asn Val Glu Val Asp
245 250 255
Thr Ala Pro Thr Val Asp Leu Asn Gln Val Leu Asn Glu Thr Arg Ser
260 265 270
Gln Tyr Glu Ala Leu Val Glu Ile Asn Arg Arg Glu Val Glu Gln Trp
275 280 285
Phe Ala Thr Gln Thr Glu Glu Leu Asn Lys Gln Val Val Ser Ser Ser
290 295 300
Glu Gln Leu Gln Ser Cys Gln Ala Glu Ile Ile Glu Leu Arg Arg Thr
305 310 315 320
Val Asn Ala Leu Glu Ile Glu Leu Gln Ala Gln His Asn Leu Arg Asp
325 330 335
Ser Leu Glu Asn Thr Leu Thr Glu Ser Glu Ala His Tyr Ser Ser Gln
340 345 350
Leu Ser Gln Val Gln Ser Leu Ile Thr Asn Val Glu Ser Gln Leu Ala
355 360 365
Glu Ile Arg Cys Asp Leu Glu Arg Gln Asn Gln Glu Tyr Gln Val Leu
370 375 380
Leu Asp Val Arg Ala Arg Leu Glu Cys Glu Ile Asn Thr Tyr Arg Ser
385 390 395 400
Leu Leu Glu Ser Glu Asp Cys Lys Leu Pro Cys Asn Pro Cys Ala Thr
405 410 415
Thr Asn Ala Ser Gly Asn Ser Cys Gly Pro Cys Gly Thr Ser Gln Lys
420 425 430
Gly Cys Cys Asn
435

<210> 428

<211> 416

<212> PRT

<213> Homo sapiens

<400> 428
Met Pro Tyr Asn Phe Cys Leu Pro Ser Leu Ser Cys Arg Thr Ser Cys
1 5 10 15
Ser Ser Arg Pro Cys Val Pro Pro Ser Cys His Ser Cys Thr Leu Pro
20 25 30
Gly Ala Cys Asn Ile Pro Ala Asn Val Ser Asn Cys Asn Trp Phe Cys
35 40 45
Glu Gly Ser Phe Asn Gly Ser Glu Lys Glu Thr Met Gln Phe Leu Asn
50 55 60
Asp Arg Leu Ala Ser Tyr Leu Glu Lys Val Arg Gln Leu Glu Arg Asp
65 70 75 80
Asn Ala Glu Leu Glu Asn Leu Ile Arg Glu Arg Ser Gln Gln Gln Glu
85 90 95
Pro Leu Leu Cys Pro Ser Tyr Gln Ser Tyr Phe Lys Thr Ile Glu Glu
100 105 110
Leu Gln Gln Lys Ile Leu Cys Thr Lys Ser Glu Asn Ala Arg Leu Val
115 120 125
Val Gln Ile Asp Asn Ala Lys Leu Ala Ala Asp Asp Phe Arg Thr Lys
130 135 140
Tyr Gln Thr Glu Leu Ser Leu Arg Gln Leu Val Glu Ser Asp Ile Asn
145 150 155 160
Gly Leu Arg Arg Ile Leu Asp Glu Leu Thr Leu Cys Lys Ser Asp Leu
165 170 175
Glu Ala Gln Val Glu Ser Leu Lys Glu Glu Leu Leu Cys Leu Lys Ser
180 185 190
Asn His Glu Gln Glu Val Asn Thr Leu Arg Cys Gln Leu Gly Asp Arg
195 200 205
Leu Asn Val Glu Val Asp Ala Ala Pro Thr Val Asp Leu Asn Arg Val
210 215 220

Leu Asn Glu Thr Arg Ser Gln Tyr Glu Ala Leu Val Glu Thr Asn Arg
225 230 235 240
Arg Glu Val Glu Gln Trp Phe Thr Thr Gln Thr Glu Glu Leu Asn Lys
245 250 255
Gln Val Val Ser Ser Ser Glu Gln Leu Gln Ser Tyr Gln Ala Glu Ile
260 265 270
Ile Glu Leu Arg Arg Thr Val Asn Ala Leu Glu Ile Glu Leu Gln Ala
275 280 285
Gln His Asn Leu Arg Asp Ser Leu Glu Asn Thr Leu Thr Glu Ser Glu
290 295 300
Ala Arg Tyr Ser Ser Gln Leu Ser Gln Val Gln Ser Leu Ile Thr Asn
305 310 315 320
Val Glu Ser Gln Leu Ala Glu Ile Arg Ser Asp Leu Glu Arg Gln Asn
325 330 335
Gln Glu Tyr Gln Val Leu Leu Asp Val Arg Ala Arg Leu Glu Cys Glu
340 345 350
Ile Asn Thr Tyr Arg Ser Leu Leu Glu Ser Glu Asp Cys Asn Leu Pro
355 360 365
Ser Asn Pro Cys Ala Thr Thr Asn Ala Cys Ser Lys Pro Ile Gly Pro
370 375 380
Cys Leu Ser Asn Pro Cys Thr Ser Cys Val Pro Pro Ala Pro Cys Thr
385 390 395 400
Pro Cys Ala Pro Arg Pro Arg Cys Gly Pro Cys Asn Ser Phe Val Arg
405 410 415

<210> 429

<211> 201

<212> PRT

<213> Homo sapiens

<400> 429
Met Thr Ser Asp His Cys Ser Ser Leu Leu Ser Gly Gln Val Ser Glu
1 5 10 15
Ala Asn Ala Ala Ser Leu Cys Leu Leu Ala Asn Val Ala His Ala Asn
20 25 30
Arg Val Arg Val Gly Ser Thr Pro Leu Gly Arg Leu Ser Leu Cys Leu
35 40 45
Pro Pro Thr Cys His Thr Thr Cys Pro Leu Pro Gly Thr Cys His Ile
50 55 60
Pro Gly Asn Ile Gly Ile Cys Gly Ala Tyr Arg Glu Asn Thr Leu Asn
65 70 75 80
Gly His Glu Lys Glu Thr Met Gln Phe Leu Asn Asp Arg Leu Ala Asn
85 90 95
Tyr Leu Glu Lys Val Arg Gln Leu Glu Trp Asp Asn Ala Glu Leu Glu
100 105 110
Thr Lys Leu His Glu Arg Ser Lys Cys His Glu Ser Ser Val Cys Arg
115 120 125
Asn Tyr Gln Ser Tyr Phe Cys Thr Ile Gln Glu Leu Gln Gln Lys Val
130 135 140
Arg Phe Ala Val His Gln Ile Arg Gly Gln Glu Ser Ala Tyr Cys Leu
145 150 155 160
Ser Ala Lys Ser Gly Pro Pro Pro Ala Phe Ala Asn Lys Val Leu Leu
165 170 175
Val His Gly His Ala His Ala Phe Val Cys Cys Leu Gln Leu Leu Leu
180 185 190
Tyr Tyr Ser Gly Arg Val Gln Ser Leu
195 200
<210> 430

<211> 471

<212> PRT

<213> Homo sapiens

<400> 430
```
Met Thr Ser Phe Tyr Ser Thr Ser Ser Cys Pro Leu Gly Cys Thr Met
1               5                   10                  15
Ala Pro Gly Ala Arg Asn Val Phe Val Ser Pro Ile Asp Val Gly Cys
            20                  25                  30
Gln Pro Val Ala Glu Ala Asn Ala Ala Ser Met Cys Leu Leu Ala Asn
        35                  40                  45
Val Ala His Ala Asn Arg Val Arg Val Gly Ser Thr Pro Leu Gly Arg
    50                  55                  60
Pro Ser Leu Cys Leu Pro Pro Thr Ser His Thr Ala Cys Pro Leu Pro
65                  70                  75                  80
Gly Thr Cys His Ile Pro Gly Asn Ile Gly Ile Cys Gly Ala Tyr Gly
                85                  90                  95
Lys Asn Thr Leu Asn Gly His Glu Lys Glu Thr Met Lys Phe Leu Asn
            100                 105                 110
Asp Arg Leu Ala Asn Tyr Leu Glu Lys Val Arg Gln Leu Glu Gln Glu
        115                 120                 125
Asn Ala Glu Leu Glu Thr Thr Leu Leu Glu Arg Ser Lys Cys His Glu
    130                 135                 140
Ser Thr Val Cys Pro Asp Tyr Gln Ser Tyr Phe Arg Thr Ile Glu Glu
145                 150                 155                 160
Leu Gln Gln Lys Ile Leu Cys Ser Lys Ala Glu Asn Ala Arg Leu Ile
                165                 170                 175
Val Gln Ile Asp Asn Ala Lys Leu Ala Ala Asp Asp Phe Arg Ile Lys
            180                 185                 190
Leu Glu Ser Glu Arg Ser Leu His Gln Leu Val Glu Ala Asp Lys Cys
        195                 200                 205
Gly Thr Gln Lys Leu Leu Asp Asp Ala Thr Leu Ala Lys Ala Asp Leu
    210                 215                 220
Glu Ala Gln Gln Glu Ser Leu Lys Glu Glu Gln Leu Ser Leu Lys Ser
225                 230                 235                 240
Asn His Glu Gln Glu Val Lys Ile Leu Arg Ser Gln Leu Gly Glu Lys
                245                 250                 255
Phe Arg Ile Glu Leu Asp Ile Glu Pro Thr Ile Asp Leu Asn Arg Val
            260                 265                 270
Leu Gly Glu Met Arg Ala Gln Tyr Glu Ala Met Val Glu Thr Asn His
        275                 280                 285
Gln Asp Val Glu Gln Trp Phe Gln Ala Gln Ser Glu Gly Ile Ser Leu
    290                 295                 300
Gln Ala Met Ser Cys Ser Glu Glu Leu Gln Cys Cys Gln Ser Glu Ile
305                 310                 315                 320
Leu Glu Leu Arg Cys Thr Val Asn Ala Leu Glu Val Glu Arg Gln Ala
                325                 330                 335
Gln His Thr Leu Lys Asp Cys Leu Gln Asn Ser Leu Cys Glu Ala Glu
            340                 345                 350
Asp Arg Tyr Gly Thr Glu Leu Ala Gln Met Gln Ser Leu Ile Ser Asn
        355                 360                 365
Leu Glu Glu Gln Leu Ser Glu Ile Arg Ala Asp Leu Glu Arg Gln Asn
    370                 375                 380
Gln Glu Tyr Gln Val Leu Leu Asp Val Lys Ala Arg Leu Glu Asn Glu
385                 390                 395                 400
Ile Ala Thr Tyr Arg Asn Leu Thr Pro Leu Gln Ser Leu Phe His Ala
                405                 410                 415
Cys Leu Leu Tyr Phe Leu Ser Lys Leu Trp Pro Cys His Arg Trp Val
            420                 425                 430
Ser Leu Trp Pro Trp Ser Gln His Gly Glu Met Ile Leu Lys Ala Arg
        435                 440                 445
Val Arg Arg Leu Arg Leu Val Ala Leu Gly Ser Gly Val Pro Ser Pro
    450                 455                 460
Cys Pro Val Phe Leu Gln Asp
465                 470
```

<210> 431

<211> 456

<212> PRT

<213> Homo sapiens

<400> 431

```
Met Thr Ser Ser Tyr Ser Ser Ser Ser Cys Pro Leu Gly Cys Thr Met
1               5                   10                  15
Ala Pro Gly Ala Arg Asn Val Ser Val Ser Pro Ile Asp Ile Gly Cys
            20                  25                  30
Gln Pro Gly Ala Glu Ala Asn Ile Ala Pro Met Cys Leu Leu Ala Asn
        35                  40                  45
Val Ala His Ala Asn Arg Val Arg Val Gly Ser Thr Pro Leu Gly Arg
    50                  55                  60
Pro Ser Leu Cys Leu Pro Pro Thr Cys His Thr Ala Cys Pro Leu Pro
65                  70                  75                  80
Gly Thr Cys His Ile Pro Gly Asn Ile Gly Ile Cys Gly Ala Tyr Gly
                85                  90                  95
Glu Asn Thr Leu Asn Gly His Glu Lys Glu Thr Met Gln Phe Leu Asn
            100                 105                 110
Asp Arg Leu Ala Asn Tyr Leu Glu Lys Val Arg Gln Leu Glu Gln Glu
        115                 120                 125
Asn Ala Glu Leu Glu Ala Thr Leu Leu Glu Arg Ser Lys Cys His Glu
    130                 135                 140
Ser Thr Val Cys Pro Asp Tyr Gln Ser Tyr Phe His Thr Ile Glu Glu
145                 150                 155                 160
Leu Gln Gln Lys Ile Leu Cys Ser Lys Ala Glu Asn Ala Arg Leu Ile
                165                 170                 175
Val Gln Ile Asp Asn Ala Lys Leu Ala Ala Asp Asp Phe Arg Ile Lys
            180                 185                 190
Leu Glu Ser Glu Arg Ser Leu Arg Gln Leu Val Glu Ala Asp Lys Cys
        195                 200                 205
Gly Thr Gln Lys Leu Leu Asp Asp Ala Thr Leu Ala Lys Ala Asp Leu
    210                 215                 220
Glu Ala Gln Gln Glu Ser Leu Lys Glu Glu Gln Leu Ser Leu Lys Ser
225                 230                 235                 240
Asn His Glu Gln Glu Val Lys Ile Leu Arg Ser Gln Leu Gly Glu Lys
                245                 250                 255
Leu Arg Ile Glu Leu Asp Ile Glu Pro Thr Ile Asp Leu Asn Arg Val
            260                 265                 270
Leu Gly Glu Met Arg Ala Gln Tyr Glu Ala Met Leu Glu Thr Asn Arg
        275                 280                 285
Gln Asp Val Glu Gln Trp Phe Gln Ala Gln Ser Glu Gly Ile Ser Leu
    290                 295                 300
Gln Asp Met Ser Cys Ser Glu Glu Leu Gln Cys Cys Gln Ser Glu Ile
305                 310                 315                 320
Leu Glu Leu Arg Cys Thr Val Asn Ala Leu Glu Val Glu Arg Gln Ala
                325                 330                 335
Gln His Thr Leu Lys Asp Cys Leu Gln Asn Ser Leu Cys Glu Ala Glu
            340                 345                 350
Asp Arg Phe Gly Thr Glu Leu Ala Gln Met Gln Ser Leu Ile Ser Asn
        355                 360                 365
Val Glu Glu Gln Leu Ser Glu Ile Arg Ala Asp Leu Glu Arg Gln Asn
    370                 375                 380
Gln Glu Tyr Gln Val Leu Leu Asp Val Lys Thr Arg Leu Glu Asn Glu
385                 390                 395                 400
Ile Ala Thr Tyr Arg Asn Leu Leu Glu Ser Glu Asp Cys Lys Leu Pro
                405                 410                 415
Cys Asn Pro Cys Ser Thr Ser Pro Ser Cys Val Thr Ala Pro Cys Ala
            420                 425                 430
```

```
Pro Arg Pro Ser Cys Gly Pro Cys Thr Thr Cys Gly Pro Thr Cys Gly
        435                 440                 445
Ala Ser Thr Thr Gly Ser Arg Phe
    450                 455
```

<210> 432

<211> 448

<212> PRT

<213> Homo sapiens

<400> 432

```
Met Thr Ser Ser Cys Cys Val Thr Asn Asn Leu Gln Ala Ser Leu Lys
1               5                   10                  15
Ser Cys Pro Arg Pro Ala Ser Val Cys Ser Ser Gly Val Asn Cys Arg
            20                  25                  30
Pro Glu Leu Cys Leu Gly Tyr Val Cys Gln Pro Met Ala Cys Leu Pro
        35                  40                  45
Ser Val Cys Leu Pro Thr Thr Phe Arg Pro Ala Ser Cys Leu Ser Lys
    50                  55                  60
Thr Tyr Leu Ser Ser Ser Cys Gln Ala Ala Ser Gly Ile Ser Gly Ser
65                  70                  75                  80
Met Gly Pro Gly Ser Trp Tyr Ser Glu Gly Ala Phe Asn Gly Asn Glu
                85                  90                  95
Lys Glu Thr Met Gln Phe Leu Asn Asp Arg Leu Ala Ser Tyr Leu Thr
            100                 105                 110
Arg Val Arg Gln Leu Glu Gln Glu Asn Ala Glu Leu Glu Ser Arg Ile
        115                 120                 125
Gln Glu Ala Ser His Ser Gln Val Leu Thr Met Thr Pro Asp Tyr Gln
    130                 135                 140
Ser His Phe Arg Thr Ile Glu Glu Leu Gln Gln Lys Ile Leu Cys Thr
145                 150                 155                 160
Lys Ala Glu Asn Ala Arg Met Val Val Asn Ile Asp Asn Ala Lys Leu
                165                 170                 175
Ala Ala Asp Asp Phe Arg Ala Lys Tyr Glu Ala Glu Leu Ala Met Arg
            180                 185                 190
Gln Leu Val Glu Ala Asp Ile Asn Gly Leu Arg Arg Ile Leu Asp Asp
        195                 200                 205
Leu Thr Leu Cys Lys Ala Asp Leu Glu Ala Gln Val Glu Ser Leu Lys
    210                 215                 220
Glu Glu Leu Met Cys Leu Lys Lys Asn His Glu Glu Glu Val Gly Ser
225                 230                 235                 240
Leu Arg Cys Gln Leu Gly Asp Arg Leu Asn Ile Glu Val Asp Ala Ala
                245                 250                 255
Pro Pro Val Asp Leu Thr Arg Val Leu Glu Glu Met Arg Cys Gln Tyr
            260                 265                 270
Glu Ala Met Val Glu Ala Asn Arg Arg Asp Val Glu Glu Trp Phe Asn
        275                 280                 285
Met Gln Met Glu Glu Leu Asn Gln Gln Val Ala Thr Ser Ser Glu Gln
    290                 295                 300
Leu Gln Asn Tyr Gln Ser Asp Ile Ile Asp Leu Arg Arg Thr Val Asn
305                 310                 315                 320
Thr Leu Glu Ile Glu Leu Gln Ala Gln His Ser Leu Arg Asp Ser Leu
                325                 330                 335
Glu Asn Thr Leu Thr Glu Ser Glu Ala Arg Tyr Ser Ser Gln Leu Ala
            340                 345                 350
Gln Met Gln Cys Met Ile Thr Asn Val Glu Ala Gln Leu Ala Glu Ile
        355                 360                 365
Arg Ala Asp Leu Glu Arg Gln Asn Gln Glu Tyr Gln Val Leu Leu Asp
    370                 375                 380
Val Arg Ala Arg Leu Glu Gly Glu Ile Asn Thr Tyr Arg Ser Leu Leu
385                 390                 395                 400
```

Glu Ser Glu Asp Cys Lys Leu Pro Cys Asn Pro Cys Ser Thr Pro Ser
405 410 415
Cys Thr Thr Cys Val Pro Ser Pro Cys Val Thr Arg Thr Val Cys Val
420 425 430
Pro Arg Thr Val Gly Met Pro Cys Ser Pro Cys Pro Gln Gly Arg Tyr
435 440 445

<210> 433

<211> 425

<212> PRT

<213> Homo sapiens

<400> 433
Met Tyr Ser Ser Ser Ser Cys Lys Leu Pro Ser Leu Ser Pro Val Ala
1 5 10 15
Arg Ser Phe Ser Ala Cys Ser Val Gly Leu Gly Arg Ser Ser Tyr Arg
20 25 30
Ala Thr Ser Cys Leu Pro Ala Leu Cys Leu Pro Ala Gly Gly Phe Ala
35 40 45
Thr Ser Tyr Ser Gly Gly Gly Gly Trp Phe Gly Glu Gly Ile Leu Thr
50 55 60
Gly Asn Glu Lys Glu Thr Met Gln Ser Leu Asn Asp Arg Leu Ala Gly
65 70 75 80
Tyr Leu Glu Lys Val Arg Gln Leu Glu Gln Glu Asn Ala Ser Leu Glu
85 90 95
Ser Arg Ile Arg Glu Trp Cys Glu Gln Gln Val Pro Tyr Met Cys Pro
100 105 110
Asp Tyr Gln Ser Tyr Phe Arg Thr Ile Glu Glu Leu Gln Lys Lys Thr
115 120 125
Leu Cys Ser Lys Ala Glu Asn Ala Arg Leu Val Val Glu Ile Asp Asn
130 135 140
Ala Lys Leu Ala Ala Asp Asp Phe Arg Thr Lys Tyr Glu Thr Glu Val
145 150 155 160
Ser Leu Arg Gln Leu Val Glu Ser Asp Ile Asn Gly Leu Arg Arg Ile
165 170 175
Leu Asp Asp Leu Thr Leu Cys Lys Ser Asp Leu Glu Ala Gln Val Glu
180 185 190
Ser Leu Lys Glu Glu Leu Leu Cys Leu Lys Lys Asn His Glu Glu Glu
195 200 205
Val Asn Ser Leu Arg Cys Gln Leu Gly Asp Arg Leu Asn Val Glu Val
210 215 220
Asp Ala Ala Pro Pro Val Asp Leu Asn Arg Val Leu Glu Glu Met Arg
225 230 235 240
Cys Gln Tyr Glu Thr Leu Val Glu Asn Asn Arg Arg Asp Ala Glu Asp
245 250 255
Trp Leu Asp Thr Gln Ser Glu Glu Leu Asn Gln Gln Val Val Ser Ser
260 265 270
Ser Glu Gln Leu Gln Ser Cys Gln Ala Glu Ile Ile Glu Leu Arg Arg
275 280 285
Thr Val Asn Ala Leu Glu Ile Glu Leu Gln Ala Gln His Ser Met Arg
290 295 300
Asp Ala Leu Glu Ser Thr Leu Ala Glu Thr Glu Ala Arg Tyr Ser Ser
305 310 315 320
Gln Leu Ala Gln Met Gln Cys Met Ile Thr Asn Val Glu Ala Gln Leu
325 330 335
Ala Glu Ile Arg Ala Asp Leu Glu Arg Gln Asn Gln Glu Tyr Gln Val
340 345 350
Leu Leu Asp Val Arg Ala Arg Leu Glu Cys Glu Ile Asn Thr Tyr Arg
355 360 365
Gly Leu Leu Glu Ser Glu Asp Ser Lys Leu Pro Cys Asn Pro Cys Ala
370 375 380

Pro Asp Tyr Ser Pro Ser Lys Ser Cys Leu Pro Cys Leu Pro Ala Ala
385 390 395 400
Ser Cys Gly Pro Ser Ala Ala Arg Thr Asn Cys Ser Pro Arg Pro Ile
405 410 415
Cys Val Pro Cys Pro Gly Gly Arg Phe
420 425

<210> 434

<211> 467

<212> PRT

<213> Homo sapiens

<400> 434
Met Ala Thr Gln Thr Cys Thr Pro Thr Phe Ser Thr Gly Ser Ile Lys
1 5 10 15
Gly Leu Cys Gly Thr Ala Gly Gly Ile Ser Arg Val Ser Ser Ile Arg
20 25 30
Ser Val Gly Ser Cys Arg Val Pro Ser Leu Ala Gly Ala Ala Gly Tyr
35 40 45
Ile Ser Ser Ala Arg Ser Gly Leu Ser Gly Leu Gly Ser Cys Leu Pro
50 55 60
Gly Ser Tyr Leu Ser Ser Glu Cys His Thr Ser Gly Phe Val Gly Ser
65 70 75 80
Gly Gly Trp Phe Cys Glu Gly Ser Phe Asn Gly Ser Glu Lys Glu Thr
85 90 95
Met Gln Phe Leu Asn Asp Arg Leu Ala Asn Tyr Leu Glu Lys Val Arg
100 105 110
Gln Leu Glu Arg Glu Asn Ala Glu Leu Glu Ser Arg Ile Gln Glu Trp
115 120 125
Tyr Glu Phe Gln Ile Pro Tyr Ile Cys Pro Asp Tyr Gln Ser Tyr Phe
130 135 140
Lys Thr Ile Glu Asp Phe Gln Gln Lys Ile Leu Leu Thr Lys Ser Glu
145 150 155 160
Asn Ala Arg Leu Val Leu Gln Ile Asp Asn Ala Lys Leu Ala Ala Asp
165 170 175
Asp Phe Arg Thr Lys Tyr Glu Thr Glu Leu Ser Leu Arg Gln Leu Val
180 185 190
Glu Ala Asp Ile Asn Gly Leu Arg Arg Ile Leu Asp Glu Leu Thr Leu
195 200 205
Cys Lys Ala Asp Leu Glu Ala Gln Val Glu Ser Leu Lys Glu Glu Leu
210 215 220
Met Cys Leu Lys Lys Asn His Glu Glu Glu Val Ser Val Leu Arg Cys
225 230 235 240
Gln Leu Gly Asp Arg Leu Asn Val Glu Val Asp Ala Ala Pro Pro Val
245 250 255
Asp Leu Asn Lys Ile Leu Glu Asp Met Arg Cys Gln Tyr Glu Ala Leu
260 265 270
Val Glu Asn Asn Arg Arg Asp Val Glu Ala Trp Phe Asn Thr Gln Thr
275 280 285
Glu Glu Leu Asn Gln Gln Val Val Ser Ser Ser Glu Gln Leu Gln Cys
290 295 300
Cys Gln Thr Glu Ile Ile Glu Leu Arg Arg Thr Val Asn Ala Leu Glu
305 310 315 320
Ile Glu Leu Gln Ala Gln His Ser Met Arg Asn Ser Leu Glu Ser Thr
325 330 335
Leu Ala Glu Thr Glu Ala Arg Tyr Ser Ser Gln Leu Ala Gln Met Gln
340 345 350
Cys Leu Ile Ser Asn Val Glu Ala Gln Leu Ser Glu Ile Arg Cys Asp
355 360 365
Leu Glu Arg Gln Asn Gln Glu Tyr Gln Val Leu Leu Asp Val Lys Ala
370 375 380

Arg Leu Glu Gly Glu Ile Ala Thr Tyr Arg His Leu Leu Glu Gly Glu
385 390 395 400
Asp Cys Lys Leu Pro Pro Gln Pro Cys Ala Thr Ala Cys Lys Pro Val
405 410 415
Ile Arg Val Pro Ser Val Pro Pro Val Pro Cys Val Pro Ser Val Pro
420 425 430
Cys Thr Pro Ala Pro Gln Val Gly Thr Gln Ile Arg Thr Ile Thr Glu
435 440 445
Glu Ile Arg Asp Gly Lys Val Ile Ser Ser Arg Glu His Val Gln Ser
450 455 460
Arg Pro Leu
465

<210> 435

<211> 420

<212> PRT

<213> Homo sapiens

<400> 435
Met Ser Leu Arg Leu Gln Ser Ser Ser Ala Ser Tyr Gly Gly Gly Phe
1 5 10 15
Gly Gly Gly Ser Cys Gln Leu Gly Gly Gly Arg Gly Val Ser Thr Cys
20 25 30
Ser Thr Arg Phe Val Ser Gly Gly Ser Ala Gly Gly Tyr Gly Gly Gly
35 40 45
Val Ser Cys Gly Phe Gly Gly Gly Ala Gly Ser Gly Phe Gly Gly Gly
50 55 60
Tyr Gly Gly Gly Leu Gly Gly Gly Tyr Gly Gly Gly Leu Gly Gly Gly
65 70 75 80
Phe Gly Gly Gly Phe Ala Gly Gly Phe Val Asp Phe Gly Ala Cys Asp
85 90 95
Gly Gly Leu Leu Thr Gly Asn Glu Lys Ile Thr Met Gln Asn Leu Asn
100 105 110
Asp Arg Leu Ala Ser Tyr Leu Glu Lys Val Arg Ala Leu Glu Glu Ala
115 120 125
Asn Ala Asp Leu Glu Val Lys Ile Arg Asp Trp His Leu Lys Gln Ser
130 135 140
Pro Ala Ser Pro Glu Arg Asp Tyr Ser Pro Tyr Tyr Lys Thr Ile Glu
145 150 155 160
Glu Leu Arg Asp Lys Ile Leu Thr Ala Thr Ile Glu Asn Asn Arg Val
165 170 175
Ile Leu Glu Ile Asp Asn Ala Arg Leu Ala Val Asp Asp Phe Arg Leu
180 185 190
Lys Tyr Glu Asn Glu Leu Ala Leu Arg Gln Ser Val Glu Ala Asp Ile
195 200 205
Asn Gly Leu Arg Arg Val Leu Asp Glu Leu Thr Leu Ser Lys Thr Asp
210 215 220
Leu Glu Met Gln Ile Glu Ser Leu Asn Glu Glu Leu Ala Tyr Met Lys
225 230 235 240
Lys Asn His Glu Glu Glu Met Lys Glu Phe Ser Asn Gln Val Val Gly
245 250 255
Gln Val Asn Val Glu Met Asp Ala Thr Pro Gly Ile Asp Leu Thr Arg
260 265 270
Val Leu Ala Glu Met Arg Glu Gln Tyr Glu Ala Met Ala Glu Arg Asn
275 280 285
Arg Arg Asp Ala Glu Glu Trp Phe His Ala Lys Ser Ala Glu Leu Asn
290 295 300
Lys Glu Val Ser Thr Asn Thr Ala Met Ile Gln Thr Ser Lys Thr Glu
305 310 315 320
Ile Thr Glu Leu Arg Arg Thr Leu Gln Gly Leu Glu Ile Glu Leu Gln
325 330 335

Ser Gln Leu Ser Met Lys Ala Gly Leu Glu Asn Thr Val Ala Glu Thr
340 345 350
Glu Cys Arg Tyr Ala Leu Gln Leu Gln Gln Ile Gln Gly Leu Ile Ser
355 360 365
Ser Ile Glu Ala Gln Leu Ser Glu Leu Arg Ser Glu Met Glu Cys Gln
370 375 380
Asn Gln Glu Tyr Lys Met Leu Leu Asp Ile Lys Thr Arg Leu Glu Gln
385 390 395 400
Glu Ile Ala Thr Tyr Arg Ser Leu Leu Glu Gly Gln Asp Ala Lys Lys
405 410 415
Arg Gln Pro Pro
420

<210> 436

<211> 456

<212> PRT

<213> Homo sapiens

<400> 436
Met Thr Thr Thr Phe Leu Gln Thr Ser Ser Ser Thr Phe Gly Gly Gly
1 5 10 15
Ser Thr Arg Gly Gly Ser Leu Leu Ala Gly Gly Gly Gly Phe Gly Gly
20 25 30
Gly Ser Leu Ser Gly Gly Gly Gly Ser Arg Ser Ile Ser Ala Ser Ser
35 40 45
Ala Arg Phe Val Ser Ser Gly Ser Gly Gly Gly Tyr Gly Gly Gly Met
50 55 60
Arg Val Cys Gly Phe Gly Gly Gly Ala Gly Ser Val Phe Gly Gly Gly
65 70 75 80
Phe Gly Gly Gly Val Gly Gly Gly Phe Gly Gly Gly Phe Gly Gly Gly
85 90 95
Asp Gly Gly Leu Leu Ser Gly Asn Glu Lys Ile Thr Met Gln Asn Leu
100 105 110
Asn Asp Arg Leu Ala Ser Tyr Leu Asp Lys Val Arg Ala Leu Glu Glu
115 120 125
Ala Asn Ala Asp Leu Glu Val Lys Ile His Asp Trp Tyr Gln Lys Gln
130 135 140
Thr Pro Thr Ser Pro Glu Cys Asp Tyr Ser Gln Tyr Phe Lys Thr Ile
145 150 155 160
Glu Glu Leu Arg Asp Lys Ile Met Ala Thr Thr Ile Asp Asn Ser Arg
165 170 175
Val Ile Leu Glu Ile Asp Asn Ala Arg Leu Ala Ala Asp Asp Phe Arg
180 185 190
Leu Lys Tyr Glu Asn Glu Leu Ala Leu Arg Gln Gly Val Glu Ala Asp
195 200 205
Ile Asn Gly Leu Arg Arg Val Leu Asp Glu Leu Thr Leu Ala Arg Thr
210 215 220
Asp Leu Glu Met Gln Ile Glu Gly Leu Asn Glu Glu Leu Ala Tyr Leu
225 230 235 240
Lys Lys Asn His Glu Glu Glu Met Lys Glu Phe Ser Ser Gln Leu Ala
245 250 255
Gly Gln Val Asn Val Glu Met Asp Ala Ala Pro Gly Val Asp Leu Thr
260 265 270
Arg Val Leu Ala Glu Met Arg Glu Gln Tyr Glu Ala Met Ala Glu Lys
275 280 285
Asn Arg Arg Asp Val Glu Ala Trp Phe Phe Ser Lys Thr Glu Glu Leu
290 295 300
Asn Lys Glu Val Ala Ser Asn Thr Glu Met Ile Gln Thr Ser Lys Thr
305 310 315 320
Glu Ile Thr Asp Leu Arg Arg Thr Met Gln Glu Leu Glu Ile Glu Leu
325 330 335

```
Gln Ser Gln Leu Ser Met Lys Ala Gly Leu Glu Asn Ser Leu Ala Glu
            340                 345                 350
Thr Glu Cys Arg Tyr Ala Thr Gln Leu Gln Gln Ile Gln Gly Leu Ile
        355                 360                 365
Gly Gly Leu Glu Ala Gln Leu Ser Glu Leu Arg Cys Glu Met Glu Ala
    370                 375                 380
Gln Asn Gln Glu Tyr Lys Met Leu Leu Asp Ile Lys Thr Arg Leu Glu
385                 390                 395                 400
Gln Glu Ile Ala Thr Tyr Arg Ser Leu Leu Glu Gly Gln Asp Ala Lys
                405                 410                 415
Met Ala Gly Ile Gly Ile Arg Glu Ala Ser Ser Gly Gly Gly Gly Ser
            420                 425                 430
Ser Ser Asn Phe His Ile Asn Val Glu Glu Ser Val Asp Gly Gln Val
        435                 440                 445
Val Ser Ser His Lys Arg Glu Ile
    450                 455
<210>  437

<211>  400

<212>  PRT

<213>  Homo sapiens


<400>  437
Met Thr Ser Tyr Ser Tyr Arg Gln Ser Ser Ala Thr Ser Ser Phe Gly
1               5                   10                  15
Gly Leu Gly Gly Gly Ser Val Arg Phe Gly Pro Gly Val Ala Phe Arg
            20                  25                  30
Ala Pro Ser Ile His Gly Gly Ser Gly Gly Arg Gly Val Ser Val Ser
        35                  40                  45
Ser Ala Arg Phe Val Ser Ser Ser Ser Ser Gly Ala Tyr Gly Gly Gly
    50                  55                  60
Tyr Gly Gly Val Leu Thr Ala Ser Asp Gly Leu Leu Ala Gly Asn Glu
65                  70                  75                  80
Lys Leu Thr Met Gln Asn Leu Asn Asp Arg Leu Ala Ser Tyr Leu Asp
                85                  90                  95
Lys Val Arg Ala Leu Glu Ala Ala Asn Gly Glu Leu Glu Val Lys Ile
            100                 105                 110
Arg Asp Trp Tyr Gln Lys Gln Gly Pro Gly Pro Ser Arg Asp Tyr Ser
        115                 120                 125
His Tyr Tyr Thr Thr Ile Gln Asp Leu Arg Asp Lys Ile Leu Gly Ala
    130                 135                 140
Thr Ile Glu Asn Ser Arg Ile Val Leu Gln Ile Asp Asn Ala Arg Leu
145                 150                 155                 160
Ala Ala Asp Asp Phe Arg Thr Lys Phe Glu Thr Glu Gln Ala Leu Arg
                165                 170                 175
Met Ser Val Glu Ala Asp Ile Asn Gly Leu Arg Arg Val Leu Asp Glu
            180                 185                 190
Leu Thr Leu Ala Arg Thr Asp Leu Glu Met Gln Ile Glu Gly Leu Lys
        195                 200                 205
Glu Glu Leu Ala Tyr Leu Lys Lys Asn His Glu Glu Glu Ile Ser Thr
    210                 215                 220
Leu Arg Gly Gln Val Gly Gly Gln Val Ser Val Glu Val Asp Ser Ala
225                 230                 235                 240
Pro Gly Thr Asp Leu Ala Lys Ile Leu Ser Asp Met Arg Ser Gln Tyr
                245                 250                 255
Glu Val Met Ala Glu Gln Asn Arg Lys Asp Ala Glu Ala Trp Phe Thr
            260                 265                 270
Ser Arg Thr Glu Glu Leu Asn Arg Glu Val Ala Gly His Thr Glu Gln
        275                 280                 285
Leu Gln Met Ser Arg Ser Glu Val Thr Asp Leu Arg Arg Thr Leu Gln
    290                 295                 300
```

Gly Leu Glu Ile Glu Leu Gln Ser Gln Leu Ser Met Lys Ala Ala Leu
305 310 315 320
Glu Asp Thr Leu Ala Glu Thr Glu Ala Arg Phe Gly Ala Gln Leu Ala
325 330 335
His Ile Gln Ala Leu Ile Ser Gly Ile Glu Ala Gln Leu Gly Asp Val
340 345 350
Arg Ala Asp Ser Glu Arg Gln Asn Gln Glu Tyr Gln Arg Leu Met Asp
355 360 365
Ile Lys Ser Arg Leu Glu Gln Glu Ile Ala Thr Tyr Arg Ser Leu Leu
370 375 380
Glu Gly Gln Glu Asp His Tyr Asn Asn Leu Ser Ala Ser Lys Val Leu
385 390 395 400

<210> 438

<211> 622

<212> PRT

<213> Homo sapiens

<400> 438
Met Ser Cys Arg Gln Phe Ser Ser Ser Tyr Leu Thr Ser Gly Gly Gly
1 5 10 15
Gly Gly Gly Gly Leu Gly Ser Gly Gly Ser Ile Arg Ser Ser Tyr Ser
20 25 30
Arg Phe Ser Ser Ser Gly Gly Arg Gly Gly Gly Gly Arg Phe Ser Ser
35 40 45
Ser Ser Gly Tyr Gly Gly Gly Ser Ser Arg Val Cys Gly Arg Gly Gly
50 55 60
Gly Gly Ser Phe Gly Tyr Ser Tyr Gly Gly Gly Ser Gly Gly Gly Phe
65 70 75 80
Ser Ala Ser Ser Leu Gly Gly Gly Phe Gly Gly Gly Ser Arg Gly Phe
85 90 95
Gly Gly Ala Ser Gly Gly Gly Tyr Ser Ser Ser Gly Gly Phe Gly Gly
100 105 110
Gly Phe Gly Gly Gly Ser Gly Gly Gly Phe Gly Gly Gly Tyr Gly Ser
115 120 125
Gly Phe Gly Gly Leu Gly Gly Phe Gly Gly Gly Ala Gly Gly Gly Asp
130 135 140
Gly Gly Ile Leu Thr Ala Asn Glu Lys Ser Thr Met Gln Glu Leu Asn
145 150 155 160
Ser Arg Leu Ala Ser Tyr Leu Asp Lys Val Gln Ala Leu Glu Glu Ala
165 170 175
Asn Asn Asp Leu Glu Asn Lys Ile Gln Asp Trp Tyr Asp Lys Lys Gly
180 185 190
Pro Ala Ala Ile Gln Lys Asn Tyr Ser Pro Tyr Tyr Asn Thr Ile Asp
195 200 205
Asp Leu Lys Asp Gln Ile Val Asp Leu Thr Val Gly Asn Asn Lys Thr
210 215 220
Leu Leu Asp Ile Asp Asn Thr Arg Met Thr Leu Asp Asp Phe Arg Ile
225 230 235 240
Lys Phe Glu Met Glu Gln Asn Leu Arg Gln Gly Val Asp Ala Asp Ile
245 250 255
Asn Gly Leu Arg Gln Val Leu Asp Asn Leu Thr Met Glu Lys Ser Asp
260 265 270
Leu Glu Met Gln Tyr Glu Thr Leu Gln Glu Glu Leu Met Ala Leu Lys
275 280 285
Lys Asn His Lys Glu Glu Met Ser Gln Leu Thr Gly Gln Asn Ser Gly
290 295 300
Asp Val Asn Val Glu Ile Asn Val Ala Pro Gly Lys Asp Leu Thr Lys
305 310 315 320
Thr Leu Asn Asp Met Arg Gln Glu Tyr Glu Gln Leu Ile Ala Lys Asn
325 330 335

Arg Lys Asp Ile Glu Asn Gln Tyr Glu Thr Gln Ile Thr Gln Ile Glu
340 345 350
His Glu Val Ser Ser Ser Gly Gln Glu Val Gln Ser Ser Ala Lys Glu
355 360 365
Val Thr Gln Leu Arg His Gly Val Gln Glu Leu Glu Ile Glu Leu Gln
370 375 380
Ser Gln Leu Ser Lys Lys Ala Ala Leu Glu Lys Ser Leu Glu Asp Thr
385 390 395 400
Lys Asn Arg Tyr Cys Gly Gln Leu Gln Met Ile Gln Glu Gln Ile Ser
405 410 415
Asn Leu Glu Ala Gln Ile Thr Asp Val Arg Gln Glu Ile Glu Cys Gln
420 425 430
Asn Gln Glu Tyr Ser Leu Leu Leu Ser Ile Lys Met Arg Leu Glu Lys
435 440 445
Glu Ile Glu Thr Tyr His Asn Leu Leu Glu Gly Gly Gln Glu Asp Phe
450 455 460
Glu Ser Ser Gly Ala Gly Lys Ile Gly Leu Gly Gly Arg Gly Gly Ser
465 470 475 480
Gly Gly Ser Tyr Gly Arg Gly Ser Arg Gly Gly Ser Gly Gly Ser Tyr
485 490 495
Gly Gly Gly Gly Ser Gly Gly Gly Tyr Gly Gly Gly Ser Gly Ser Arg
500 505 510
Gly Gly Ser Gly Gly Ser Tyr Gly Gly Gly Ser Gly Ser Gly Gly Gly
515 520 525
Ser Gly Gly Gly Tyr Gly Gly Gly Ser Gly Gly Gly His Ser Gly Gly
530 535 540
Ser Gly Gly Gly His Ser Gly Gly Ser Gly Gly Asn Tyr Gly Gly Gly
545 550 555 560
Ser Gly Ser Gly Gly Gly Ser Gly Gly Gly Tyr Gly Gly Gly Ser Gly
565 570 575
Ser Arg Gly Gly Ser Gly Gly Ser His Gly Gly Gly Ser Gly Phe Gly
580 585 590
Gly Glu Ser Gly Gly Ser Tyr Gly Gly Gly Glu Glu Ala Ser Gly Ser
595 600 605
Gly Gly Gly Tyr Gly Gly Gly Ser Gly Lys Ser Ser His Ser
610 615 620

<210> 439

<211> 472

<212> PRT

<213> Homo sapiens

<400> 439
Met Thr Thr Cys Ser Arg Gln Phe Thr Ser Ser Ser Ser Met Lys Gly
1 5 10 15
Ser Cys Gly Ile Gly Gly Gly Ile Gly Gly Gly Ser Ser Arg Ile Ser
20 25 30
Ser Val Leu Ala Gly Gly Ser Cys Arg Ala Pro Ser Thr Tyr Gly Gly
35 40 45
Gly Leu Ser Val Ser Ser Ser Arg Phe Ser Ser Gly Gly Ala Tyr Gly
50 55 60
Leu Gly Gly Gly Tyr Gly Gly Gly Phe Ser Ser Ser Ser Ser Ser Phe
65 70 75 80
Gly Ser Gly Phe Gly Gly Gly Tyr Gly Gly Gly Leu Gly Ala Gly Leu
85 90 95
Gly Gly Gly Phe Gly Gly Gly Phe Ala Gly Gly Asp Gly Leu Leu Val
100 105 110
Gly Ser Glu Lys Val Thr Met Gln Asn Leu Asn Asp Arg Leu Ala Ser
115 120 125
Tyr Leu Asp Lys Val Arg Ala Leu Glu Glu Ala Asn Ala Asp Leu Glu
130 135 140

Val Lys Ile Arg Asp Trp Tyr Gln Arg Gln Arg Pro Ala Glu Ile Lys
145 150 155 160
Asp Tyr Ser Pro Tyr Phe Lys Thr Ile Glu Asp Leu Arg Asn Lys Ile
165 170 175
Leu Thr Ala Thr Val Asp Asn Ala Asn Val Leu Leu Gln Ile Asp Asn
180 185 190
Ala Arg Leu Ala Ala Asp Asp Phe Arg Thr Lys Tyr Glu Thr Glu Leu
195 200 205
Asn Leu Arg Met Ser Val Glu Ala Asp Ile Asn Gly Leu Arg Arg Val
210 215 220
Leu Asp Glu Leu Thr Leu Ala Arg Ala Asp Leu Glu Met Gln Ile Glu
225 230 235 240
Ser Leu Lys Glu Glu Leu Ala Tyr Leu Lys Lys Asn His Glu Glu Glu
245 250 255
Met Asn Ala Leu Arg Gly Gln Val Gly Gly Asp Val Asn Val Glu Met
260 265 270
Asp Ala Ala Pro Gly Val Asp Leu Ser Arg Ile Leu Asn Glu Met Arg
275 280 285
Asp Gln Tyr Glu Lys Met Ala Glu Lys Asn Arg Lys Asp Ala Glu Glu
290 295 300
Trp Phe Phe Thr Lys Thr Glu Glu Leu Asn Arg Glu Val Ala Thr Asn
305 310 315 320
Ser Glu Leu Val Gln Ser Gly Lys Ser Glu Ile Ser Glu Leu Arg Arg
325 330 335
Thr Met Gln Asn Leu Glu Ile Glu Leu Gln Ser Gln Leu Ser Met Lys
340 345 350
Ala Ser Leu Glu Asn Ser Leu Glu Glu Thr Lys Gly Arg Tyr Cys Met
355 360 365
Gln Leu Ala Gln Ile Gln Glu Met Ile Gly Ser Val Glu Glu Gln Leu
370 375 380
Ala Gln Leu Arg Cys Glu Met Glu Gln Gln Asn Gln Glu Tyr Lys Ile
385 390 395 400
Leu Leu Asp Val Lys Thr Arg Leu Glu Gln Glu Ile Ala Thr Tyr Arg
405 410 415
Arg Leu Leu Glu Gly Glu Asp Ala His Leu Ser Ser Ser Gln Phe Ser
420 425 430
Ser Gly Ser Gln Ser Ser Arg Asp Val Thr Ser Ser Ser Arg Gln Ile
435 440 445
Arg Thr Lys Val Met Asp Val His Asp Gly Lys Val Val Ser Thr His
450 455 460
Glu Gln Val Leu Arg Thr Lys Asn
465 470

<210> 440

<211> 473

<212> PRT

<213> Homo sapiens

<400> 440
Met Thr Thr Cys Ser Arg Gln Phe Thr Ser Ser Ser Ser Met Lys Gly
1 5 10 15
Ser Cys Gly Ile Gly Gly Gly Ile Gly Gly Gly Ser Ser Arg Ile Ser
20 25 30
Ser Val Leu Ala Gly Gly Ser Cys Arg Ala Pro Ser Thr Tyr Gly Gly
35 40 45
Gly Leu Ser Val Ser Ser Arg Phe Ser Ser Gly Gly Ala Cys Gly Leu
50 55 60
Gly Gly Gly Tyr Gly Gly Gly Phe Ser Ser Ser Ser Ser Phe Gly Ser
65 70 75 80
Gly Phe Gly Gly Gly Tyr Gly Gly Gly Leu Gly Ala Gly Phe Gly Gly
85 90 95

Gly Leu Gly Ala Gly Phe Gly Gly Gly Phe Ala Gly Gly Asp Gly Leu
100 105 110
Leu Val Gly Ser Glu Lys Val Thr Met Gln Asn Leu Asn Asp Arg Leu
115 120 125
Ala Ser Tyr Leu Asp Lys Val Arg Ala Leu Glu Glu Ala Asn Ala Asp
130 135 140
Leu Glu Val Lys Ile Arg Asp Trp Tyr Gln Arg Gln Arg Pro Ser Glu
145 150 155 160
Ile Lys Asp Tyr Ser Pro Tyr Phe Lys Thr Ile Glu Asp Leu Arg Asn
165 170 175
Lys Ile Ile Ala Ala Thr Ile Glu Asn Ala Gln Pro Ile Leu Gln Ile
180 185 190
Asp Asn Ala Arg Leu Ala Ala Asp Asp Phe Arg Thr Lys Tyr Glu His
195 200 205
Glu Leu Ala Leu Arg Gln Thr Val Glu Ala Asp Val Asn Gly Leu Arg
210 215 220
Arg Val Leu Asp Glu Leu Thr Leu Ala Arg Thr Asp Leu Glu Met Gln
225 230 235 240
Ile Glu Gly Leu Lys Glu Glu Leu Ala Tyr Leu Arg Lys Asn His Glu
245 250 255
Glu Glu Met Leu Ala Leu Arg Gly Gln Thr Gly Gly Asp Val Asn Val
260 265 270
Glu Met Asp Ala Ala Pro Gly Val Asp Leu Ser Arg Ile Leu Asn Glu
275 280 285
Met Arg Asp Gln Tyr Glu Gln Met Ala Glu Lys Asn Arg Arg Asp Ala
290 295 300
Glu Thr Trp Phe Leu Ser Lys Thr Glu Glu Leu Asn Lys Glu Val Ala
305 310 315 320
Ser Asn Ser Glu Leu Val Gln Ser Ser Arg Ser Glu Val Thr Glu Leu
325 330 335
Arg Arg Val Leu Gln Gly Leu Glu Ile Glu Leu Gln Ser Gln Leu Ser
340 345 350
Met Lys Ala Ser Leu Glu Asn Ser Leu Glu Glu Thr Lys Gly Arg Tyr
355 360 365
Cys Met Gln Leu Ser Gln Ile Gln Gly Leu Ile Gly Ser Val Glu Glu
370 375 380
Gln Leu Ala Gln Leu Arg Cys Glu Met Glu Gln Gln Ser Gln Glu Tyr
385 390 395 400
Gln Ile Leu Leu Asp Val Lys Thr Arg Leu Glu Gln Glu Ile Ala Thr
405 410 415
Tyr Arg Arg Leu Leu Glu Gly Glu Asp Ala His Leu Ser Ser Gln Gln
420 425 430
Ala Ser Gly Gln Ser Tyr Ser Ser Arg Glu Val Phe Thr Ser Ser Ser
435 440 445
Ser Ser Ser Ser Arg Gln Thr Arg Pro Ile Leu Lys Glu Gln Ser Ser
450 455 460
Ser Ser Phe Ser Gln Gly Gln Ser Ser
465 470

<210> 441

<211> 432

<212> PRT

<213> Homo sapiens

<400> 441
Met Thr Thr Ser Ile Arg Gln Phe Thr Ser Ser Ser Ser Ile Lys Gly
1 5 10 15
Ser Ser Gly Leu Gly Gly Gly Ser Ser Arg Thr Ser Cys Arg Leu Ser
20 25 30
Gly Gly Leu Gly Ala Gly Ser Cys Arg Leu Gly Ser Ala Gly Gly Leu
35 40 45

Gly Ser Thr Leu Gly Gly Ser Ser Tyr Ser Ser Cys Tyr Ser Phe Gly
50 55 60
Ser Gly Gly Gly Tyr Gly Ser Ser Phe Gly Gly Val Asp Gly Leu Leu
65 70 75 80
Ala Gly Gly Glu Lys Ala Thr Met Gln Asn Leu Asn Asp Arg Leu Ala
85 90 95
Ser Tyr Leu Asp Lys Val Arg Ala Leu Glu Glu Ala Asn Thr Glu Leu
100 105 110
Glu Val Lys Ile Arg Asp Trp Tyr Gln Arg Gln Ala Pro Gly Pro Ala
115 120 125
Arg Asp Tyr Ser Gln Tyr Tyr Arg Thr Ile Glu Glu Leu Gln Asn Lys
130 135 140
Ile Leu Thr Ala Thr Val Asp Asn Ala Asn Ile Leu Leu Gln Ile Asp
145 150 155 160
Asn Ala Arg Leu Ala Ala Asp Asp Phe Arg Thr Lys Phe Glu Thr Glu
165 170 175
Gln Ala Leu Arg Leu Ser Val Glu Ala Asp Ile Asn Gly Leu Arg Arg
180 185 190
Val Leu Asp Glu Leu Thr Leu Ala Arg Ala Asp Leu Glu Met Gln Ile
195 200 205
Glu Asn Leu Lys Glu Glu Leu Ala Tyr Leu Lys Lys Asn His Glu Glu
210 215 220
Glu Met Asn Ala Leu Arg Gly Gln Val Gly Gly Glu Ile Asn Val Glu
225 230 235 240
Met Asp Ala Ala Pro Gly Val Asp Leu Ser Arg Ile Leu Asn Glu Met
245 250 255
Arg Asp Gln Tyr Glu Lys Met Ala Glu Lys Asn Arg Lys Asp Ala Glu
260 265 270
Asp Trp Phe Phe Ser Lys Thr Glu Glu Leu Asn Arg Glu Val Ala Thr
275 280 285
Asn Ser Glu Leu Val Gln Ser Gly Lys Ser Glu Ile Ser Glu Leu Arg
290 295 300
Arg Thr Met Gln Ala Leu Glu Ile Glu Leu Gln Ser Gln Leu Ser Met
305 310 315 320
Lys Ala Ser Leu Glu Gly Asn Leu Ala Glu Thr Glu Asn Arg Tyr Cys
325 330 335
Val Gln Leu Ser Gln Ile Gln Gly Leu Ile Gly Ser Val Glu Glu Gln
340 345 350
Leu Ala Gln Leu Arg Cys Glu Met Glu Gln Gln Asn Gln Glu Tyr Lys
355 360 365
Ile Leu Leu Asp Val Lys Thr Arg Leu Glu Gln Glu Ile Ala Thr Tyr
370 375 380
Arg Arg Leu Leu Glu Gly Glu Asp Ala His Leu Thr Gln Tyr Lys Lys
385 390 395 400
Glu Pro Val Thr Thr Arg Gln Val Arg Thr Ile Val Glu Glu Val Gln
405 410 415
Asp Gly Lys Val Ile Ser Ser Arg Glu Gln, Val His Gln Thr Thr Arg
420 425 430

<210> 442

<211> 469

<212> PRT

<213> Homo sapiens

<400> 442
Met Ser Ile His Phe Ser Ser Pro Val Phe Thr Ser Arg Ser Ala Ala
1 5 10 15
Phe Ser Gly Arg Gly Ala Gln Val Arg Leu Ser Ser Ala Arg Pro Gly
20 25 30
Gly Leu Gly Ser Ser Ser Leu Tyr Gly Leu Gly Ala Ser Arg Pro Arg
35 40 45

Val Ala Val Arg Ser Ala Tyr Gly Gly Pro Val Gly Ala Gly Ile Arg
50 55 60
Glu Val Thr Ile Asn Gln Ser Leu Leu Ala Pro Leu Arg Leu Asp Ala
65 70 75 80
Asp Pro Ser Leu Gln Arg Val Arg Gln Glu Glu Ser Glu Gln Ile Lys
85 90 95
Thr Leu Asn Asn Lys Phe Ala Ser Phe Ile Asp Lys Val Arg Phe Leu
100 105 110
Glu Gln Gln Asn Lys Leu Leu Glu Thr Lys Trp Thr Leu Leu Gln Glu
115 120 125
Gln Lys Ser Ala Lys Ser Ser Arg Leu Pro Asp Ile Phe Glu Ala Gln
130 135 140
Ile Ala Gly Leu Arg Gly Gln Leu Glu Ala Leu Gln Val Asp Gly Gly
145 150 155 160
Arg Leu Glu Ala Glu Leu Arg Ser Met Gln Asp Val Val Glu Asp Phe
165 170 175
Lys Asn Lys Tyr Glu Asp Glu Ile Asn Arg Arg Thr Ala Ala Glu Asn
180 185 190
Glu Phe Val Val Leu Lys Lys Asp Val Asp Ala Ala Tyr Met Ser Lys
195 200 205
Val Glu Leu Glu Ala Lys Val Asp Ala Leu Asn Asp Glu Ile Asn Phe
210 215 220
Leu Arg Thr Leu Asn Glu Thr Glu Leu Thr Glu Leu Gln Ser Gln Ile
225 230 235 240
Ser Asp Thr Ser Val Val Leu Ser Met Asp Asn Ser Arg Ser Leu Asp
245 250 255
Leu Asp Gly Ile Ile Ala Glu Val Lys Ala Gln Tyr Glu Glu Met Ala
260 265 270
Lys Cys Ser Arg Ala Glu Ala Glu Ala Trp Tyr Gln Thr Lys Phe Glu
275 280 285
Thr Leu Gln Ala Gln Ala Gly Lys His Gly Asp Asp Leu Arg Asn Thr
290 295 300
Arg Asn Glu Ile Ser Glu Met Asn Arg Ala Ile Gln Arg Leu Gln Ala
305 310 315 320
Glu Ile Asp Asn Ile Lys Asn Gln Arg Ala Lys Leu Glu Ala Ala Ile
325 330 335
Ala Glu Ala Glu Glu Arg Gly Glu Leu Ala Leu Lys Asp Ala Arg Ala
340 345 350
Lys Gln Glu Glu Leu Glu Ala Ala Leu Gln Arg Ala Lys Gln Asp Met
355 360 365
Ala Arg Gln Leu Arg Glu Tyr Gln Glu Leu Met Ser Val Lys Leu Ala
370 375 380
Leu Asp Ile Glu Ile Ala Thr Tyr Arg Lys Leu Leu Glu Gly Glu Glu
385 390 395 400
Ser Arg Leu Ala Gly Asp Gly Val Gly Ala Val Asn Ile Ser Val Met
405 410 415
Asn Ser Thr Gly Gly Ser Ser Ser Gly Gly Gly Ile Gly Leu Thr Leu
420 425 430
Gly Gly Thr Met Gly Ser Asn Ala Leu Ser Phe Ser Ser Ser Ala Gly
435 440 445
Pro Gly Leu Leu Lys Ala Tyr Ser Ile Arg Thr Ala Ser Ala Ser Arg
450 455 460
Arg Ser Ala Arg Asp
465

<210> 443

<211> 486

<212> PRT

<213> Homo sapiens

<400> 443
Met Thr Cys Gly Ser Tyr Cys Gly Gly Arg Ala Phe Ser Cys Ile Ser
1 5 10 15
Ala Cys Gly Pro Arg Pro Gly Arg Cys Cys Ile Thr Ala Ala Pro Tyr
20 25 30
Arg Gly Ile Ser Cys Tyr Arg Gly Leu Thr Gly Gly Phe Gly Ser His
35 40 45
Ser Val Cys Gly Gly Phe Arg Ala Gly Ser Cys Gly Arg Ser Phe Gly
50 55 60
Tyr Arg Ser Gly Gly Val Cys Gly Pro Ser Pro Pro Cys Ile Thr Thr
65 70 75 80
Val Ser Val Asn Glu Ser Leu Leu Thr Pro Leu Asn Leu Glu Ile Asp
85 90 95
Pro Asn Ala Gln Cys Val Lys Gln Glu Glu Lys Glu Gln Ile Lys Ser
100 105 110
Leu Asn Ser Arg Phe Ala Ala Phe Ile Asp Lys Val Arg Phe Leu Glu
115 120 125
Gln Gln Asn Lys Leu Leu Glu Thr Lys Leu Gln Phe Tyr Gln Asn Arg
130 135 140
Glu Cys Cys Gln Ser Asn Leu Glu Pro Leu Phe Glu Gly Tyr Ile Glu
145 150 155 160
Thr Leu Arg Arg Glu Ala Glu Cys Val Glu Ala Asp Ser Gly Arg Leu
165 170 175
Ala Ser Glu Leu Asn His Val Gln Glu Val Leu Glu Gly Tyr Lys Lys
180 185 190
Lys Tyr Glu Glu Glu Val Ser Leu Arg Ala Thr Ala Glu Asn Glu Phe
195 200 205
Val Ala Leu Lys Lys Asp Val Asp Cys Ala Tyr Leu Arg Lys Ser Asp
210 215 220
Leu Glu Ala Asn Val Glu Ala Leu Ile Gln Glu Ile Asp Phe Leu Arg
225 230 235 240
Arg Leu Tyr Glu Glu Glu Ile Arg Val Leu Gln Ser His Ile Ser Asp
245 250 255
Thr Ser Val Val Val Lys Leu Asp Asn Ser Arg Asp Leu Asn Met Asp
260 265 270
Cys Ile Ile Ala Glu Ile Lys Ala Gln Tyr Asp Asp Ile Val Thr Arg
275 280 285
Ser Arg Ala Glu Ala Glu Ser Trp Tyr Arg Ser Lys Cys Glu Glu Met
290 295 300
Lys Ala Thr Val Ile Arg His Gly Glu Thr Leu Arg Arg Thr Lys Glu
305 310 315 320
Glu Ile Asn Glu Leu Asn Arg Met Ile Gln Arg Leu Thr Ala Glu Val
325 330 335
Glu Asn Ala Lys Cys Gln Asn Ser Lys Leu Glu Ala Ala Val Ala Gln
340 345 350
Ser Glu Gln Gln Gly Glu Ala Ala Leu Ser Asp Ala Arg Cys Lys Leu
355 360 365
Ala Glu Leu Glu Gly Ala Leu Gln Lys Ala Lys Gln Asp Met Ala Cys
370 375 380
Leu Ile Arg Glu Tyr Gln Glu Val Met Asn Ser Lys Leu Gly Leu Asp
385 390 395 400
Ile Glu Ile Ala Thr Tyr Arg Arg Leu Leu Glu Gly Glu Glu Gln Arg
405 410 415
Leu Cys Glu Gly Val Gly Ser Val Asn Val Cys Val Ser Ser Ser Arg
420 425 430
Gly Gly Val Val Cys Gly Asp Leu Cys Ala Ser Thr Thr Ala Pro Val
435 440 445
Val Ser Thr Arg Val Ser Ser Val Pro Ser Asn Ser Asn Val Val Val
450 455 460
Gly Thr Thr Asn Ala Cys Ala Pro Ser Ala Arg Val Gly Val Cys Gly
465 470 475 480
Gly Ser Cys Lys Arg Cys
485

<210> 444

<211> 111

<212> PRT

<213> Homo sapiens

<400> 444

Met Lys Ala Thr Val Ile Trp His Gly Glu Thr Val Gly Cys Thr Lys
1 5 10 15
Glu Glu Ile Lys Glu Leu Thr His Met Ile Gln Arg Leu Met Ala Lys
20 25 30
Val Glu Asn Ala Lys Cys Gln Val Trp Gly Ile Cys Ala Gln Gly Gln
35 40 45
Arg Asp Leu Trp Pro Asn Leu Cys His Thr Ala Tyr Val Cys Pro Thr
50 55 60
Trp Ile Ser Ala Phe Ile Leu Gln Ser Leu Cys Pro Cys Arg Val Pro
65 70 75 80
Gly Cys Gly Gln Ser Gly Ser Ala Arg Met Met Lys Ala Arg Gly Leu
85 90 95
Phe Leu Arg Cys Ser Gln Leu Asn Gly Arg Leu Asp Ile Phe Arg
100 105 110

<210> 445

<211> 505

<212> PRT

<213> Homo sapiens

<400> 445

Met Thr Cys Gly Ser Gly Phe Gly Gly Arg Ala Phe Ser Cys Ile Ser
1 5 10 15
Ala Cys Gly Pro Arg Pro Gly Arg Cys Cys Ile Thr Ala Ala Pro Tyr
20 25 30
Arg Gly Ile Ser Cys Tyr Arg Gly Leu Thr Gly Gly Phe Gly Ser His
35 40 45
Ser Val Cys Gly Gly Phe Arg Ala Gly Ser Cys Gly Arg Ser Phe Gly
50 55 60
Tyr Arg Ser Gly Gly Val Cys Gly Pro Ser Pro Pro Cys Ile Thr Thr
65 70 75 80
Val Ser Val Asn Glu Ser Leu Leu Thr Pro Leu Asn Leu Glu Ile Asp
85 90 95
Pro Asn Ala Gln Cys Val Lys Gln Glu Glu Lys Glu Gln Ile Lys Ser
100 105 110
Leu Asn Ser Arg Phe Ala Ala Phe Ile Asp Lys Val Arg Phe Leu Glu
115 120 125
Gln Gln Asn Lys Leu Leu Glu Thr Lys Leu Gln Phe Tyr Gln Asn Arg
130 135 140
Glu Cys Cys Gln Ser Asn Leu Glu Pro Leu Phe Glu Gly Tyr Ile Glu
145 150 155 160
Thr Leu Arg Arg Glu Ala Glu Cys Val Glu Ala Asp Ser Gly Arg Leu
165 170 175
Ala Ser Glu Leu Asn His Val Gln Glu Val Leu Glu Gly Tyr Lys Lys
180 185 190
Lys Tyr Glu Glu Glu Val Ser Leu Arg Ala Thr Ala Glu Asn Glu Phe
195 200 205
Val Ala Leu Lys Lys Asp Val Asp Cys Ala Tyr Leu Arg Lys Ser Asp
210 215 220
Leu Glu Ala Asn Val Glu Ala Leu Ile Gln Glu Ile Asp Phe Leu Arg
225 230 235 240
Arg Leu Tyr Glu Glu Glu Ile Arg Ile Leu Gln Ser His Ile Ser Asp
245 250 255

```
Thr Ser Val Val Val Lys Leu Asp Asn Ser Arg Asp Leu Asn Met Asp
            260                 265                 270
Cys Ile Ile Ala Glu Ile Lys Ala Gln Tyr Asp Asp Ile Val Thr Arg
        275                 280                 285
Ser Arg Ala Glu Ala Glu Ser Trp Tyr Arg Ser Lys Cys Glu Glu Met
    290                 295                 300
Lys Ala Thr Val Ile Arg His Gly Glu Thr Leu Arg Arg Thr Lys Glu
305                 310                 315                 320
Glu Ile Asn Glu Leu Asn Arg Met Ile Gln Arg Leu Thr Ala Glu Val
                325                 330                 335
Glu Asn Ala Lys Cys Gln Asn Ser Lys Leu Glu Ala Ala Val Ala Gln
            340                 345                 350
Ser Glu Gln Gln Gly Glu Ala Ala Leu Ser Asp Ala Arg Cys Lys Leu
        355                 360                 365
Ala Glu Leu Glu Gly Ala Leu Gln Lys Ala Lys Gln Asp Met Ala Cys
    370                 375                 380
Leu Ile Arg Glu Tyr Gln Glu Val Met Asn Ser Lys Leu Gly Leu Asp
385                 390                 395                 400
Ile Glu Ile Ala Thr Tyr Arg Arg Leu Leu Glu Gly Glu Glu Gln Arg
                405                 410                 415
Leu Cys Glu Gly Ile Gly Ala Val Asn Val Cys Val Ser Ser Ser Arg
            420                 425                 430
Gly Gly Val Val Cys Gly Asp Leu Cys Val Ser Gly Ser Arg Pro Val
        435                 440                 445
Thr Gly Ser Val Cys Ser Ala Pro Cys Asn Gly Asn Val Ala Val Ser
    450                 455                 460
Thr Gly Leu Cys Ala Pro Cys Gly Gln Leu Asn Thr Thr Cys Gly Gly
465                 470                 475                 480
Gly Ser Cys Gly Val Gly Ser Cys Gly Ile Ser Ser Leu Gly Val Gly
                485                 490                 495
Ser Cys Gly Ser Ser Cys Arg Lys Cys
            500                 505
```

<210> 446

<211> 486

<212> PRT

<213> Homo sapiens

<400> 446

```
Met Thr Cys Gly Ser Tyr Cys Gly Gly Arg Ala Phe Ser Cys Ile Ser
1               5                   10                  15
Ala Cys Gly Pro Arg Pro Gly Arg Cys Cys Ile Thr Ala Ala Pro Tyr
            20                  25                  30
Arg Gly Ile Ser Cys Tyr Arg Gly Leu Thr Gly Gly Phe Gly Ser His
        35                  40                  45
Ser Val Cys Gly Gly Phe Arg Ala Gly Ser Cys Gly Arg Ser Phe Gly
    50                  55                  60
Tyr Arg Ser Gly Gly Val Cys Gly Pro Ser Pro Pro Cys Ile Thr Thr
65                  70                  75                  80
Val Ser Val Asn Glu Ser Leu Leu Thr Pro Leu Asn Leu Glu Ile Asp
                85                  90                  95
Pro Asn Ala Gln Cys Val Lys Gln Glu Glu Lys Glu Gln Ile Lys Ser
            100                 105                 110
Leu Asn Ser Arg Phe Ala Ala Phe Ile Asp Lys Val Arg Phe Leu Glu
        115                 120                 125
Gln Gln Asn Lys Leu Leu Glu Thr Lys Leu Gln Phe Tyr Gln Asn Arg
    130                 135                 140
Glu Cys Cys Gln Ser Asn Leu Glu Pro Leu Phe Glu Gly Tyr Ile Glu
145                 150                 155                 160
Thr Leu Arg Arg Glu Ala Glu Cys Val Glu Ala Asp Ser Gly Arg Leu
                165                 170                 175
```

Ala Ser Glu Leu Asn His Val Gln Glu Val Leu Glu Gly Tyr Lys Lys
180 185 190
Lys Tyr Glu Glu Glu Val Ser Leu Arg Ala Thr Ala Glu Asn Glu Phe
195 200 205
Val Ala Leu Lys Lys Asp Val Asp Cys Ala Tyr Leu Arg Lys Ser Asp
210 215 220
Leu Glu Ala Asn Val Glu Ala Leu Ile Gln Glu Ile Asp Phe Leu Arg
225 230 235 240
Arg Leu Tyr Glu Glu Glu Ile Arg Val Leu Gln Ser His Ile Ser Asp
245 250 255
Thr Ser Val Val Val Lys Leu Asp Asn Ser Arg Asp Leu Asn Met Asp
260 265 270
Cys Ile Ile Ala Glu Ile Lys Ala Gln Tyr Asp Asp Ile Val Thr Arg
275 280 285
Ser Arg Ala Glu Ala Glu Ser Trp Tyr Arg Ser Lys Cys Glu Glu Met
290 295 300
Lys Ala Thr Val Ile Arg His Gly Glu Thr Leu Arg Arg Thr Lys Glu
305 310 315 320
Glu Ile Asn Glu Leu Asn Arg Met Ile Gln Arg Leu Thr Ala Glu Val
325 330 335
Glu Asn Ala Lys Cys Gln Asn Ser Lys Leu Glu Ala Ala Val Ala Gln
340 345 350
Ser Glu Gln Gln Gly Glu Ala Ala Leu Ser Asp Ala Arg Cys Lys Leu
355 360 365
Ala Glu Leu Glu Gly Ala Leu Gln Lys Ala Lys Gln Asp Met Ala Cys
370 375 380
Leu Ile Arg Glu Tyr Gln Glu Val Met Asn Ser Lys Leu Gly Leu Asp
385 390 395 400
Ile Glu Ile Ala Thr Tyr Arg Arg Leu Leu Glu Gly Glu Glu Gln Arg
405 410 415
Leu Cys Glu Gly Val Gly Ser Val Asn Val Cys Val Ser Ser Ser Arg
420 425 430
Gly Gly Val Val Cys Gly Asp Leu Cys Ala Ser Thr Thr Ala Pro Val
435 440 445
Val Ser Thr Arg Val Ser Ser Val Pro Ser Asn Ser Asn Val Val Val
450 455 460
Gly Thr Thr Asn Ala Cys Ala Pro Ser Ala Arg Val Gly Val Cys Gly
465 470 475 480
Gly Ser Cys Lys Arg Cys
485

<210> 447

<211> 493

<212> PRT

<213> Homo sapiens

<400> 447
Met Thr Cys Gly Phe Asn Ser Ile Gly Cys Gly Phe Arg Pro Gly Asn
1 5 10 15
Phe Ser Cys Val Ser Ala Cys Gly Pro Arg Pro Ser Arg Cys Cys Ile
20 25 30
Thr Ala Ala Pro Tyr Arg Gly Ile Ser Cys Tyr Arg Gly Leu Thr Gly
35 40 45
Gly Phe Gly Ser His Ser Val Cys Gly Gly Phe Arg Ala Gly Ser Cys
50 55 60
Gly Arg Ser Phe Gly Tyr Arg Ser Gly Gly Val Cys Gly Pro Ser Pro
65 70 75 80
Pro Cys Ile Thr Thr Val Ser Val Asn Glu Ser Leu Leu Thr Pro Leu
85 90 95
Asn Leu Glu Ile Asp Pro Asn Ala Gln Cys Val Lys Gln Glu Glu Lys
100 105 110

Glu Gln Ile Lys Ser Leu Asn Ser Arg Phe Ala Ala Phe Ile Asp Lys
115 120 125
Val Arg Phe Leu Glu Gln Gln Asn Lys Leu Leu Glu Thr Lys Leu Gln
130 135 140
Phe Tyr Gln Asn Cys Glu Cys Cys Gln Ser Asn Leu Glu Pro Leu Phe
145 150 155 160
Ala Gly Tyr Ile Glu Thr Leu Arg Arg Glu Ala Glu Cys Val Glu Ala
165 170 175
Asp Ser Gly Arg Leu Ala Ser Glu Leu Asn His Val Gln Glu Val Leu
180 185 190
Glu Gly Tyr Lys Lys Lys Tyr Glu Glu Glu Val Ala Leu Arg Ala Thr
195 200 205
Ala Glu Asn Glu Phe Val Ala Leu Lys Lys Asp Val Asp Cys Ala Tyr
210 215 220
Leu Arg Lys Ser Asp Leu Glu Ala Asn Val Glu Ala Leu Ile Gln Glu
225 230 235 240
Ile Asp Phe Leu Arg Arg Leu Tyr Glu Glu Glu Ile Arg Ile Leu Gln
245 250 255
Ser His Ile Ser Asp Thr Ser Val Val Val Lys Leu Asp Asn Ser Arg
260 265 270
Asp Leu Asn Met Asp Cys Ile Val Ala Glu Ile Lys Ala Gln Tyr Asp
275 280 285
Asp Ile Ala Thr Arg Ser Arg Ala Glu Ala Glu Ser Trp Tyr Arg Ser
290 295 300
Lys Cys Glu Glu Met Lys Ala Thr Val Ile Arg His Gly Glu Thr Leu
305 310 315 320
Arg Arg Thr Lys Glu Glu Ile Asn Glu Leu Asn Arg Met Ile Gln Arg
325 330 335
Leu Thr Ala Glu Val Glu Asn Ala Lys Cys Gln Asn Ser Lys Leu Glu
340 345 350
Ala Ala Val Ala Gln Ser Glu Gln Gln Gly Glu Ala Ala Leu Ser Asp
355 360 365
Ala Arg Cys Lys Leu Ala Glu Leu Glu Gly Ala Leu Gln Lys Ala Lys
370 375 380
Gln Asp Met Ala Cys Leu Ile Arg Glu Tyr Gln Glu Val Met Asn Ser
385 390 395 400
Lys Leu Gly Leu Asp Ile Glu Ile Ala Thr Tyr Arg Arg Leu Leu Glu
405 410 415
Gly Glu Glu Gln Arg Leu Cys Glu Gly Val Glu Ala Val Asn Val Cys
420 425 430
Val Ser Ser Ser Arg Gly Gly Val Val Cys Gly Asp Leu Cys Val Ser
435 440 445
Gly Ser Arg Pro Val Thr Gly Ser Val Cys Ser Ala Pro Cys Asn Gly
450 455 460
Asn Leu Val Val Ser Thr Gly Leu Cys Lys Pro Cys Gly Gln Leu Asn
465 470 475 480
Thr Thr Cys Gly Gly Gly Ser Cys Gly Gln Gly Arg Tyr
485 490

<210> 448

<211> 143

<212> PRT

<213> Homo sapiens

<400> 448
Met Ala Ser Gln Ser Cys His Ile Ser Ser Gly Cys Gly Val Lys Asn
1 5 10 15
Phe Ser Ser Arg Ser Ala Thr Val Pro Lys Pro Gly Tyr His Ser Cys
20 25 30
Val Ser Ala Met Ala His His Gly Val Ser Pro Gly Gly Leu Gly Ser
35 40 45

Arg Arg Leu Gly Gly Phe Gly Ser Gln Ser Leu Cys Thr Val Gly Ser
50 55 60
Pro Arg Ile Ala Val Ser Cys Arg Trp Pro Leu His Ser Arg Gly Arg
65 70 75 80
Phe Gly Tyr Trp Ala Gly Gly Leu Cys Arg Pro Ser Pro Pro Arg Ile
85 90 95
Thr Ser Val Thr Ile Asn Glu Ser Leu Leu Met Pro Leu Asn Leu Glu
100 105 110
Ile Asp Pro Asn Ala Gln Cys Val Lys His Glu Glu Lys Glu His Ile
115 120 125
Arg Cys Leu Asn Lys Phe Ala Ala Phe Ile Asp Lys Val Gly Leu
130 135 140

<210> 449

<211> 507

<212> PRT

<213> Homo sapiens

<400> 449

Met Ser Cys Arg Ser Tyr Arg Ile Ser Ser Gly Cys Gly Val Thr Arg
1 5 10 15
Asn Phe Ser Ser Cys Ser Ala Val Ala Pro Lys Thr Gly Asn Arg Cys
20 25 30
Cys Ile Ser Ala Ala Pro Tyr Arg Gly Val Ser Cys Tyr Arg Gly Leu
35 40 45
Thr Gly Phe Gly Ser Arg Ser Leu Cys Asn Leu Gly Ser Cys Gly Pro
50 55 60
Arg Ile Ala Val Gly Gly Phe Arg Ala Gly Ser Cys Gly Arg Ser Phe
65 70 75 80
Gly Tyr Arg Ser Gly Gly Val Cys Gly Pro Ser Pro Pro Cys Ile Thr
85 90 95
Thr Val Ser Val Asn Glu Ser Leu Leu Thr Pro Leu Asn Leu Glu Ile
100 105 110
Asp Pro Asn Ala Gln Cys Val Lys Gln Glu Glu Lys Glu Gln Ile Lys
115 120 125
Ser Leu Asn Ser Arg Phe Ala Ala Phe Ile Asp Lys Val Arg Phe Leu
130 135 140
Glu Gln Gln Asn Lys Leu Leu Glu Thr Lys Trp Gln Phe Tyr Gln Asn
145 150 155 160
Gln Arg Cys Cys Glu Ser Asn Leu Glu Pro Leu Phe Ser Gly Tyr Ile
165 170 175
Glu Thr Leu Arg Arg Glu Ala Glu Cys Val Glu Ala Asp Ser Gly Arg
180 185 190
Leu Ala Ser Glu Leu Asn His Val Gln Glu Val Leu Glu Gly Tyr Lys
195 200 205
Lys Lys Tyr Glu Glu Glu Val Ala Leu Arg Ala Thr Ala Glu Asn Glu
210 215 220
Phe Val Val Leu Lys Lys Asp Val Asp Cys Ala Tyr Leu Arg Lys Ser
225 230 235 240
Asp Leu Glu Ala Asn Val Glu Ala Leu Val Glu Glu Ser Ser Phe Leu
245 250 255
Arg Arg Leu Tyr Glu Glu Glu Ile Arg Val Leu Gln Ala His Ile Ser
260 265 270
Asp Thr Ser Val Ile Val Lys Met Asp Asn Ser Arg Asp Leu Asn Met
275 280 285
Asp Cys Ile Ile Ala Glu Ile Lys Ala Gln Tyr Asp Asp Val Ala Ser
290 295 300
Arg Ser Arg Ala Glu Ala Glu Ser Trp Tyr Arg Ser Lys Cys Glu Glu
305 310 315 320
Met Lys Ala Thr Val Ile Arg His Gly Glu Thr Leu Arg Arg Thr Lys
325 330 335

Glu Glu Ile Asn Glu Leu Asn Arg Met Ile Gln Arg Leu Thr Ala Glu
340 345 350
Ile Glu Asn Ala Lys Cys Gln Arg Ala Lys Leu Glu Ala Ala Val Ala
355 360 365
Glu Ala Glu Gln Gln Gly Glu Ala Ala Leu Ser Asp Ala Arg Cys Lys
370 375 380
Leu Ala Glu Leu Glu Gly Ala Leu Gln Lys Ala Lys Gln Asp Met Ala
385 390 395 400
Cys Leu Leu Lys Glu Tyr Gln Glu Val Met Asn Ser Lys Leu Gly Leu
405 410 415
Asp Ile Glu Ile Ala Thr Tyr Arg Arg Leu Leu Glu Gly Glu Glu His
420 425 430
Arg Leu Cys Glu Gly Val Gly Ser Val Asn Val Cys Val Ser Ser Ser
435 440 445
Arg Gly Gly Val Ser Cys Gly Gly Leu Ser Tyr Ser Thr Thr Pro Gly
450 455 460
Arg Gln Ile Thr Ser Gly Pro Ser Ala Ile Gly Gly Ser Ile Thr Val
465 470 475 480
Val Ala Pro Asp Ser Cys Ala Pro Cys Gln Pro Arg Ser Ser Ser Phe
485 490 495
Ser Cys Gly Ser Ser Arg Ser Val Arg Phe Ala
500 505

<210> 450

<211> 600

<212> PRT

<213> Homo sapiens

<400> 450
Met Ser Cys Arg Ser Tyr Arg Val Ser Ser Gly His Arg Val Gly Asn
1 5 10 15
Phe Ser Ser Cys Ser Ala Met Thr Pro Gln Asn Leu Asn Arg Phe Arg
20 25 30
Ala Asn Ser Val Ser Cys Trp Ser Gly Pro Gly Phe Arg Gly Leu Gly
35 40 45
Ser Phe Gly Ser Arg Ser Val Ile Thr Phe Gly Ser Tyr Ser Pro Arg
50 55 60
Ile Ala Ala Val Gly Ser Arg Pro Ile His Cys Gly Val Arg Phe Gly
65 70 75 80
Ala Gly Cys Gly Met Gly Phe Gly Asp Gly Arg Gly Val Gly Leu Gly
85 90 95
Pro Arg Ala Asp Ser Cys Val Gly Leu Gly Phe Gly Ala Gly Ser Gly
100 105 110
Ile Gly Tyr Gly Phe Gly Gly Pro Gly Phe Gly Tyr Arg Val Gly Gly
115 120 125
Val Gly Val Pro Ala Ala Pro Ser Ile Thr Ala Val Thr Val Asn Lys
130 135 140
Ser Leu Leu Thr Pro Leu Asn Leu Glu Ile Asp Pro Asn Ala Gln Arg
145 150 155 160
Val Lys Lys Asp Glu Lys Glu Gln Ile Lys Thr Leu Asn Asn Lys Phe
165 170 175
Ala Ser Phe Ile Asp Lys Val Arg Phe Leu Glu Gln Gln Asn Lys Leu
180 185 190
Leu Glu Thr Lys Trp Ser Phe Leu Gln Glu Gln Lys Cys Ile Arg Ser
195 200 205
Asn Leu Glu Pro Leu Phe Glu Ser Tyr Ile Thr Asn Leu Arg Arg Gln
210 215 220
Leu Glu Val Leu Val Ser Asp Gln Ala Arg Leu Gln Ala Glu Arg Asn
225 230 235 240
His Leu Gln Asp Val Leu Glu Gly Phe Lys Lys Lys Tyr Glu Glu Glu
245 250 255

Val Val Cys Arg Ala Asn Ala Glu Asn Glu Phe Val Ala Leu Lys Lys
260 265 270
Asp Val Asp Ala Ala Phe Met Asn Lys Ser Asp Leu Glu Ala Asn Val
275 280 285
Asp Thr Leu Thr Gln Glu Ile Asp Phe Leu Lys Thr Leu Tyr Met Glu
290 295 300
Glu Ile Gln Leu Leu Gln Ser His Ile Ser Glu Thr Ser Val Ile Val
305 310 315 320
Lys Met Asp Asn Ser Arg Asp Leu Asn Leu Asp Gly Ile Ile Ala Glu
325 330 335
Val Lys Ala Gln Tyr Glu Glu Val Ala Arg Arg Ser Arg Ala Asp Ala
340 345 350
Glu Ala Trp Tyr Gln Thr Lys Tyr Glu Glu Met Gln Val Thr Ala Gly
355 360 365
Gln His Cys Asp Asn Leu Arg Asn Ile Arg Asn Glu Ile Asn Glu Leu
370 375 380
Thr Arg Leu Ile Gln Arg Leu Lys Ala Glu Ile Glu His Ala Lys Ala
385 390 395 400
Gln Arg Ala Lys Leu Glu Ala Ala Val Ala Glu Ala Glu Gln Gln Gly
405 410 415
Glu Ala Thr Leu Ser Asp Ala Lys Cys Lys Leu Ala Asp Leu Glu Cys
420 425 430
Ala Leu Gln Gln Ala Lys Gln Asp Met Ala Arg Gln Leu Cys Glu Tyr
435 440 445
Gln Glu Leu Met Asn Ala Lys Leu Gly Leu Asp Ile Glu Ile Ala Thr
450 455 460
Tyr Arg Arg Leu Leu Glu Gly Glu Glu Ser Arg Leu Cys Glu Gly Val
465 470 475 480
Gly Pro Val Asn Ile Ser Val Ser Ser Ser Arg Gly Gly Leu Val Cys
485 490 495
Gly Pro Glu Pro Leu Val Ala Gly Ser Thr Leu Ser Arg Gly Gly Val
500 505 510
Thr Phe Ser Gly Ser Ser Ser Val Cys Ala Thr Ser Gly Val Leu Ala
515 520 525
Ser Cys Gly Pro Ser Leu Gly Gly Ala Arg Val Ala Pro Ala Thr Gly
530 535 540
Asp Leu Leu Ser Thr Gly Thr Arg Ser Gly Ser Met Leu Ile Ser Glu
545 550 555 560
Ala Cys Val Pro Ser Val Pro Cys Pro Leu Pro Thr Gln Gly Gly Phe
565 570 575
Ser Ser Cys Ser Gly Gly Arg Ser Ser Ser Val Arg Phe Val Ser Thr
580 585 590
Thr Thr Ser Cys Arg Thr Lys Tyr
595 600

<210> 451

<211> 513

<212> PRT

<213> Homo sapiens

<400> 451
Met Ser Tyr His Ser Phe Gln Pro Gly Ser Arg Cys Gly Ser Gln Ser
1 5 10 15
Phe Ser Ser Tyr Ser Ala Val Met Pro Arg Met Val Thr His Tyr Ala
20 25 30
Val Ser Lys Gly Pro Cys Arg Pro Gly Gly Gly Arg Gly Leu Arg Ala
35 40 45
Leu Gly Cys Leu Gly Ser Arg Ser Leu Cys Asn Val Gly Phe Gly Arg
50 55 60
Pro Arg Val Ala Ser Arg Cys Gly Gly Thr Leu Pro Gly Phe Gly Tyr
65 70 75 80

```
Arg Leu Gly Ala Thr Cys Gly Pro Ser Ala Cys Ile Thr Pro Val Thr
                85                  90                  95
Ile Asn Glu Ser Leu Leu Val Pro Leu Ala Leu Glu Ile Asp Pro Thr
            100                 105                 110
Val Gln Arg Val Lys Arg Asp Glu Lys Glu Gln Ile Lys Cys Leu Asn
        115                 120                 125
Asn Arg Phe Ala Ser Phe Ile Asn Lys Val Arg Phe Leu Glu Gln Lys
    130                 135                 140
Asn Lys Leu Leu Glu Thr Lys Trp Asn Phe Met Gln Gln Gln Arg Cys
145                 150                 155                 160
Cys Gln Thr Asn Ile Glu Pro Ile Phe Glu Gly Tyr Ile Ser Ala Leu
                165                 170                 175
Arg Arg Gln Leu Asp Cys Val Ser Gly Asp Arg Val Arg Leu Glu Ser
            180                 185                 190
Glu Leu Cys Ser Leu Gln Ala Ala Leu Glu Gly Tyr Lys Lys Lys Tyr
        195                 200                 205
Glu Glu Glu Leu Ser Leu Arg Pro Cys Val Glu Asn Glu Phe Val Ala
    210                 215                 220
Leu Lys Lys Asp Val Asp Thr Ala Phe Leu Met Lys Ala Asp Leu Glu
225                 230                 235                 240
Thr Asn Ala Glu Ala Leu Val Gln Glu Ile Asp Phe Leu Lys Ser Leu
                245                 250                 255
Tyr Glu Glu Glu Ile Cys Leu Leu Gln Ser Gln Ile Ser Glu Thr Ser
            260                 265                 270
Val Ile Val Lys Met Asp Asn Ser Arg Glu Leu Asp Val Asp Gly Ile
        275                 280                 285
Ile Ala Glu Ile Lys Ala Gln Tyr Asp Asp Ile Ala Ser Arg Ser Lys
    290                 295                 300
Ala Glu Ala Glu Ala Trp Tyr Gln Cys Arg Tyr Glu Glu Leu Arg Val
305                 310                 315                 320
Thr Ala Gly Asn His Cys Asp Asn Leu Arg Asn Arg Lys Asn Glu Ile
                325                 330                 335
Leu Glu Met Asn Lys Leu Ile Gln Arg Leu Gln Gln Glu Thr Glu Asn
            340                 345                 350
Val Lys Ala Gln Arg Cys Lys Leu Glu Gly Ala Ile Ala Glu Ala Glu
        355                 360                 365
Gln Gln Gly Glu Ala Ala Leu Asn Asp Ala Lys Cys Lys Leu Ala Gly
    370                 375                 380
Leu Glu Glu Ala Leu Gln Lys Ala Lys Gln Asp Met Ala Cys Leu Leu
385                 390                 395                 400
Lys Glu Tyr Gln Glu Val Met Asn Ser Lys Leu Gly Leu Asp Ile Glu
                405                 410                 415
Ile Ala Thr Tyr Arg Arg Leu Leu Glu Gly Glu Glu His Arg Leu Cys
            420                 425                 430
Glu Gly Ile Gly Pro Val Asn Ile Ser Val Ser Ser Ser Lys Gly Ala
        435                 440                 445
Phe Leu Tyr Glu Pro Cys Gly Val Ser Thr Pro Val Leu Ser Thr Gly
    450                 455                 460
Val Leu Arg Ser Asn Gly Gly Cys Ser Ile Val Gly Thr Gly Glu Leu
465                 470                 475                 480
Tyr Val Pro Cys Glu Pro Gln Gly Leu Leu Ser Cys Gly Ser Gly Arg
                485                 490                 495
Lys Ser Ser Met Thr Leu Gly Ala Gly Gly Ser Ser Pro Ser His Lys
            500                 505                 510
His
```

<210> 452

<211> 85

<212> PRT

<213> Homo sapiens

<400> 452
Met Asp Ala Val Tyr Met Asn Lys Val Gly Leu Glu Ala Lys Val Asp
1 5 10 15
Ala Leu Met Glu Glu Thr Asn Phe Leu Ser Thr Phe Tyr Lys Ala Val
20 25 30
Arg Val Pro Gly Ala Pro Ser Asn Arg Gly Ala Gly Gly Trp Val Leu
35 40 45
Glu Pro Gln Leu Gly Thr Glu Pro Val Gly Ser Phe Pro Gly Leu Leu
50 55 60
Ser Ala Pro Tyr Pro Thr Cys Val Leu Gln Gly Arg Cys His Phe Pro
65 70 75 80
Tyr His Arg Arg Lys
85

<210> 453

<211> 564

<212> PRT

<213> Homo sapiens

<400> 453
Met Ala Ser Thr Ser Thr Thr Ile Arg Ser His Ser Ser Ser Arg Arg
1 5 10 15
Gly Phe Ser Ala Ser Ser Ala Arg Leu Pro Gly Val Ser Arg Ser Gly
20 25 30
Phe Ser Ser Ile Ser Val Ser Arg Ser Arg Gly Ser Gly Gly Leu Gly
35 40 45
Gly Ala Cys Gly Gly Ala Gly Phe Gly Ser Arg Ser Leu Tyr Gly Leu
50 55 60
Gly Gly Ser Lys Arg Ile Ser Ile Gly Gly Gly Ser Cys Ala Ile Ser
65 70 75 80
Gly Gly Tyr Gly Ser Arg Ala Gly Gly Ser Tyr Gly Phe Gly Gly Ala
85 90 95
Gly Ser Gly Phe Gly Phe Gly Gly Gly Ala Gly Ile Gly Phe Gly Leu
100 105 110
Gly Gly Gly Ala Gly Leu Ala Gly Gly Phe Gly Gly Pro Gly Phe Pro
115 120 125
Val Cys Pro Pro Gly Gly Ile Gln Glu Val Thr Val Asn Gln Ser Leu
130 135 140
Leu Thr Pro Leu Asn Leu Gln Ile Asp Pro Ala Ile Gln Arg Val Arg
145 150 155 160
Ala Glu Glu Arg Glu Gln Ile Lys Thr Leu Asn Asn Lys Phe Ala Ser
165 170 175
Phe Ile Asp Lys Val Arg Phe Leu Glu Gln Gln Asn Lys Val Leu Asp
180 185 190
Thr Lys Trp Thr Leu Leu Gln Glu Gln Gly Thr Lys Thr Val Arg Gln
195 200 205
Asn Leu Glu Pro Leu Phe Glu Gln Tyr Ile Asn Asn Leu Arg Arg Gln
210 215 220
Leu Asp Ser Ile Val Gly Glu Arg Gly Arg Leu Asp Ser Glu Leu Arg
225 230 235 240
Asn Met Gln Asp Leu Val Glu Asp Leu Lys Asn Lys Tyr Glu Asp Glu
245 250 255
Ile Asn Lys Arg Thr Ala Ala Glu Asn Glu Phe Val Thr Leu Lys Lys
260 265 270
Asp Val Asp Ala Ala Tyr Met Asn Lys Val Glu Leu Gln Ala Lys Ala
275 280 285
Asp Thr Leu Thr Asp Glu Ile Asn Phe Leu Arg Ala Leu Tyr Asp Ala
290 295 300
Glu Leu Ser Gln Met Gln Thr His Ile Ser Asp Thr Ser Val Val Leu
305 310 315 320
Ser Met Asp Asn Asn Arg Asn Leu Asp Leu Asp Ser Ile Ile Ala Glu
325 330 335

Val Lys Ala Gln Tyr Glu Glu Ile Ala Gln Arg Ser Arg Ala Glu Ala
340 345 350
Glu Ser Trp Tyr Gln Thr Lys Tyr Glu Glu Leu Gln Val Thr Ala Gly
355 360 365
Arg His Gly Asp Asp Leu Arg Asn Thr Lys Gln Glu Ile Ala Glu Ile
370 375 380
Asn Arg Met Ile Gln Arg Leu Arg Ser Glu Ile Asp His Val Lys Lys
385 390 395 400
Gln Cys Ala Asn Leu Gln Ala Ala Ile Ala Asp Ala Glu Gln Arg Gly
405 410 415
Glu Met Ala Leu Lys Asp Ala Lys Asn Lys Leu Glu Gly Leu Glu Asp
420 425 430
Ala Leu Gln Lys Ala Lys Gln Asp Leu Ala Arg Leu Leu Lys Glu Tyr
435 440 445
Gln Glu Leu Met Asn Val Lys Leu Ala Leu Asp Val Glu Ile Ala Thr
450 455 460
Tyr Arg Lys Leu Leu Glu Gly Glu Glu Cys Arg Leu Asn Gly Glu Gly
465 470 475 480
Val Gly Gln Val Asn Ile Ser Val Val Gln Ser Thr Val Ser Ser Gly
485 490 495
Tyr Gly Gly Ala Ser Gly Val Gly Ser Gly Leu Gly Leu Gly Gly Gly
500 505 510
Ser Ser Tyr Ser Tyr Gly Ser Gly Leu Gly Val Gly Gly Gly Phe Ser
515 520 525
Ser Ser Ser Gly Arg Ala Thr Gly Gly Gly Leu Ser Ser Val Gly Gly
530 535 540
Gly Ser Ser Thr Ile Lys Tyr Thr Thr Thr Ser Ser Ser Ser Arg Lys
545 550 555 560
Ser Tyr Lys His

<210> 454

<211> 564

<212> PRT

<213> Homo sapiens

<400> 454
Met Ala Ser Thr Ser Thr Thr Ile Arg Ser His Ser Ser Ser Arg Arg
1 5 10 15
Gly Phe Ser Ala Asn Ser Ala Arg Leu Pro Gly Val Ser Arg Ser Gly
20 25 30
Phe Ser Ser Ile Ser Val Ser Arg Ser Arg Gly Ser Gly Gly Leu Gly
35 40 45
Gly Ala Cys Gly Gly Ala Gly Phe Gly Ser Arg Ser Leu Tyr Gly Leu
50 55 60
Gly Gly Ser Lys Arg Ile Ser Ile Gly Gly Gly Ser Cys Ala Ile Ser
65 70 75 80
Gly Gly Tyr Gly Ser Arg Ala Arg Gly Ser Tyr Gly Phe Gly Gly Ala
85 90 95
Gly Ser Gly Phe Gly Phe Gly Gly Gly Ala Gly Ile Gly Phe Asp Leu
100 105 110
Gly Gly Gly Ala Gly Leu Ala Gly Gly Phe Gly Gly Pro Gly Phe Pro
115 120 125
Val Cys Pro Pro Gly Gly Ile Gln Glu Val Thr Val Asn Gln Ser Leu
130 135 140
Leu Thr Pro Leu Asn Leu Gln Ile Asp Pro Ala Ile Gln Arg Val Arg
145 150 155 160
Ala Glu Glu Arg Glu Gln Ile Lys Thr Leu Asn Asn Lys Phe Ala Ser
165 170 175
Phe Ile Asp Lys Val Arg Phe Leu Glu Gln Gln Asn Lys Val Leu Asp
180 185 190

Thr Lys Trp Thr Leu Leu Gln Glu Gln Gly Thr Lys Thr Val Arg Gln
195 200 205
Asn Leu Glu Pro Leu Phe Glu Gln Tyr Ile Asn Asn Leu Arg Arg Gln
210 215 220
Leu Asp Asn Ile Val Gly Glu Arg Gly Arg Leu Asp Ser Glu Leu Arg
225 230 235 240
Asn Met Gln Asp Leu Val Glu Asp Leu Lys Asn Lys Tyr Glu Asp Glu
245 250 255
Ile Asn Lys Arg Thr Ala Ala Glu Asn Glu Phe Val Thr Leu Lys Lys
260 265 270
Asp Val Asp Ala Ala Tyr Met Asn Lys Val Glu Leu Gln Ala Lys Ala
275 280 285
Asp Thr Leu Thr Asp Glu Ile Asn Phe Leu Arg Ala Leu Tyr Asp Ala
290 295 300
Glu Leu Ser Gln Met Gln Thr His Ile Ser Asp Thr Ser Val Val Leu
305 310 315 320
Ser Met Asp Asn Asn Arg Asn Leu Asp Leu Asp Ser Ile Ile Ala Glu
325 330 335
Val Lys Ala Gln Tyr Glu Glu Ile Ala Gln Arg Ser Arg Ala Glu Ala
340 345 350
Glu Ser Trp Tyr Gln Thr Lys Tyr Glu Glu Leu Gln Val Thr Ala Gly
355 360 365
Arg His Gly Asp Asp Leu Arg Asn Thr Lys Gln Glu Ile Ala Glu Ile
370 375 380
Asn Arg Met Ile Gln Arg Leu Arg Ser Glu Ile Asp His Val Lys Lys
385 390 395 400
Gln Cys Ala Ser Leu Gln Ala Ala Ile Ala Asp Ala Glu Gln Arg Gly
405 410 415
Glu Met Ala Leu Lys Asp Ala Lys Asn Lys Leu Glu Gly Leu Glu Asp
420 425 430
Ala Leu Gln Lys Ala Lys Gln Asp Leu Ala Arg Leu Leu Lys Glu Tyr
435 440 445
Gln Glu Leu Met Asn Val Lys Leu Ala Leu Asp Val Glu Ile Ala Thr
450 455 460
Tyr Arg Lys Leu Leu Glu Gly Glu Glu Cys Arg Leu Asn Gly Glu Gly
465 470 475 480
Ile Gly Gln Val Asn Val Ser Val Val Gln Ser Thr Ile Ser Ser Gly
485 490 495
Tyr Gly Gly Ala Ser Gly Val Gly Ser Gly Leu Gly Leu Gly Gly Gly
500 505 510
Ser Ser Tyr Ser Tyr Gly Ser Gly Leu Gly Ile Gly Gly Gly Phe Ser
515 520 525
Ser Ser Ser Gly Arg Ala Ile Gly Gly Gly Leu Ser Ser Val Gly Gly
530 535 540
Gly Ser Ser Thr Ile Lys Tyr Thr Thr Thr Ser Ser Ser Ser Arg Lys
545 550 555 560
Ser Tyr Lys His

<210> 455

<211> 564

<212> PRT

<213> Homo sapiens

<400> 455
Met Ala Ser Thr Ser Thr Thr Ile Arg Ser His Ser Ser Ser Arg Arg
1 5 10 15
Gly Phe Ser Ala Asn Ser Ala Arg Leu Pro Gly Val Ser Arg Ser Gly
20 25 30
Phe Ser Ser Ile Ser Val Ser Arg Ser Arg Gly Ser Gly Gly Leu Gly
35 40 45

Gly Ala Cys Gly Gly Ala Gly Phe Gly Ser Arg Ser Leu Tyr Gly Leu
50 55 60
Gly Gly Ser Lys Arg Ile Ser Ile Gly Gly Gly Ser Cys Ala Ile Ser
65 70 75 80
Gly Gly Tyr Gly Ser Arg Ala Arg Ala Ser Tyr Gly Phe Gly Gly Ala
85 90 95
Gly Ser Gly Phe Gly Phe Gly Gly Gly Ala Gly Ile Gly Phe Asp Leu
100 105 110
Gly Gly Gly Ala Gly Leu Ala Gly Gly Phe Gly Gly Pro Gly Phe Pro
115 120 125
Val Cys Pro Pro Gly Gly Ile Gln Glu Val Thr Val Asn Gln Ser Leu
130 135 140
Leu Thr Pro Leu Asn Leu Gln Ile Asp Pro Ala Ile Gln Arg Val Arg
145 150 155 160
Ala Glu Glu Arg Glu Gln Ile Lys Thr Leu Asn Asn Lys Phe Ala Ser
165 170 175
Phe Ile Asp Lys Val Arg Phe Leu Glu Gln Gln Asn Lys Val Leu Glu
180 185 190
Thr Lys Trp Thr Leu Leu Gln Glu Gln Gly Thr Lys Thr Val Arg Gln
195 200 205
Asn Leu Glu Pro Leu Phe Glu Gln Tyr Ile Asn Asn Leu Arg Arg Gln
210 215 220
Leu Asp Ser Ile Val Gly Glu Arg Gly Arg Leu Asp Ser Glu Leu Arg
225 230 235 240
Gly Met Gln Asp Leu Val Glu Asp Phe Lys Asn Lys Tyr Glu Asp Glu
245 250 255
Ile Asn Lys Arg Thr Ala Ala Glu Asn Glu Phe Val Thr Leu Lys Lys
260 265 270
Asp Val Asp Ala Ala Tyr Met Asn Lys Val Glu Leu Gln Ala Lys Ala
275 280 285
Asp Thr Leu Thr Asp Glu Ile Asn Phe Leu Arg Ala Leu Tyr Asp Ala
290 295 300
Glu Leu Ser Gln Met Gln Thr His Ile Ser Asp Thr Ser Val Val Leu
305 310 315 320
Ser Met Asp Asn Asn Arg Asn Leu Asp Leu Asp Ser Ile Ile Ala Glu
325 330 335
Val Lys Ala Gln Tyr Glu Glu Ile Ala Gln Arg Ser Arg Ala Glu Ala
340 345 350
Glu Ser Trp Tyr Gln Thr Lys Tyr Glu Glu Leu Gln Val Thr Ala Gly
355 360 365
Arg His Gly Asp Asp Leu Arg Asn Thr Lys Gln Glu Ile Ala Glu Ile
370 375 380
Asn Arg Met Ile Gln Arg Leu Arg Ser Glu Ile Asp His Val Lys Lys
385 390 395 400
Gln Cys Ala Asn Leu Gln Ala Ala Ile Ala Asp Ala Glu Gln Arg Gly
405 410 415
Glu Met Ala Leu Lys Asp Ala Lys Asn Lys Leu Glu Gly Leu Glu Asp
420 425 430
Ala Leu Gln Lys Ala Lys Gln Asp Leu Ala Arg Leu Leu Lys Glu Tyr
435 440 445
Gln Glu Leu Met Asn Val Lys Leu Ala Leu Asp Val Glu Ile Ala Thr
450 455 460
Tyr Arg Lys Leu Leu Glu Gly Glu Glu Cys Arg Leu Asn Gly Glu Gly
465 470 475 480
Val Gly Gln Val Asn Ile Ser Val Val Gln Ser Thr Val Ser Ser Gly
485 490 495
Tyr Gly Gly Ala Ser Gly Val Gly Ser Gly Leu Gly Leu Gly Gly Gly
500 505 510
Ser Ser Tyr Ser Tyr Gly Ser Gly Leu Gly Val Gly Gly Gly Phe Ser
515 520 525
Ser Ser Ser Gly Arg Ala Ile Gly Gly Gly Leu Ser Ser Val Gly Gly
530 535 540
Gly Ser Ser Thr Ile Lys Tyr Thr Thr Thr Ser Ser Ser Ser Arg Lys
545 550 555 560
Ser Tyr Lys His

<210> 456

<211> 564

<212> PRT

<213> Homo sapiens

<400> 456

Met Ala Ser Thr Ser Thr Thr Ile Arg Ser His Ser Ser Ser Arg Arg
1 5 10 15
Gly Phe Ser Ala Asn Ser Ala Arg Leu Pro Gly Val Ser Arg Ser Gly
20 25 30
Phe Ser Ser Val Ser Val Ser Arg Ser Arg Gly Ser Gly Gly Leu Gly
35 40 45
Gly Ala Cys Gly Gly Ala Gly Phe Gly Ser Arg Ser Leu Tyr Gly Leu
50 55 60
Gly Gly Ser Lys Arg Ile Ser Ile Gly Gly Gly Ser Cys Ala Ile Ser
65 70 75 80
Gly Gly Tyr Gly Ser Arg Ala Gly Gly Ser Tyr Gly Phe Gly Gly Ala
85 90 95
Gly Ser Gly Phe Gly Phe Gly Gly Gly Ala Gly Ile Gly Phe Gly Leu
100 105 110
Gly Gly Gly Ala Gly Leu Ala Gly Gly Phe Gly Gly Pro Gly Phe Pro
115 120 125
Val Cys Pro Pro Gly Gly Ile Gln Glu Val Thr Val Asn Gln Ser Leu
130 135 140
Leu Thr Pro Leu Asn Leu Gln Ile Asp Pro Thr Ile Gln Arg Val Arg
145 150 155 160
Ala Glu Glu Arg Glu Gln Ile Lys Thr Leu Asn Asn Lys Phe Ala Ser
165 170 175
Phe Ile Asp Lys Val Arg Phe Leu Glu Gln Gln Asn Lys Val Leu Glu
180 185 190
Thr Lys Trp Thr Leu Leu Gln Glu Gln Gly Thr Lys Thr Val Arg Gln
195 200 205
Asn Leu Glu Pro Leu Phe Glu Gln Tyr Ile Asn Asn Leu Arg Arg Gln
210 215 220
Leu Asp Ser Ile Val Gly Glu Arg Gly Arg Leu Asp Ser Glu Leu Arg
225 230 235 240
Gly Met Gln Asp Leu Val Glu Asp Phe Lys Asn Lys Tyr Glu Asp Glu
245 250 255
Ile Asn Lys Arg Thr Ala Ala Glu Asn Glu Phe Val Thr Leu Lys Lys
260 265 270
Asp Val Asp Ala Ala Tyr Met Asn Lys Val Glu Leu Gln Ala Lys Ala
275 280 285
Asp Thr Leu Thr Asp Glu Ile Asn Phe Leu Arg Ala Leu Tyr Asp Ala
290 295 300
Glu Leu Ser Gln Met Gln Thr His Ile Ser Asp Thr Ser Val Val Leu
305 310 315 320
Ser Met Asp Asn Asn Arg Asn Leu Asp Leu Asp Ser Ile Ile Ala Glu
325 330 335
Val Lys Ala Gln Tyr Glu Glu Ile Ala Gln Arg Ser Arg Ala Glu Ala
340 345 350
Glu Ser Trp Tyr Gln Thr Lys Tyr Glu Glu Leu Gln Val Thr Ala Gly
355 360 365
Arg His Gly Asp Asp Leu Arg Asn Thr Lys Gln Glu Ile Ala Glu Ile
370 375 380
Asn Arg Met Ile Gln Arg Leu Arg Ser Glu Ile Asp His Val Lys Lys
385 390 395 400
Gln Cys Ala Asn Leu Gln Ala Ala Ile Ala Asp Ala Glu Gln Arg Gly
405 410 415
Glu Met Ala Leu Lys Asp Ala Lys Asn Lys Leu Glu Gly Leu Glu Asp
420 425 430

Ala Leu Gln Lys Ala Lys Gln Asp Leu Ala Arg Leu Leu Lys Glu Tyr
435 440 445
Gln Glu Leu Met Asn Val Lys Leu Ala Leu Asp Val Glu Ile Ala Thr
450 455 460
Tyr Arg Lys Leu Leu Glu Gly Glu Glu Cys Arg Leu Asn Gly Glu Gly
465 470 475 480
Val Gly Gln Val Asn Ile Ser Val Val Gln Ser Thr Val Ser Ser Gly
485 490 495
Tyr Gly Gly Ala Ser Gly Val Gly Ser Gly Leu Gly Leu Gly Gly Gly
500 505 510
Ser Ser Tyr Ser Tyr Gly Ser Gly Leu Gly Val Gly Gly Gly Phe Ser
515 520 525
Ser Ser Ser Gly Arg Ala Ile Gly Gly Gly Leu Ser Ser Val Gly Gly
530 535 540
Gly Ser Ser Thr Ile Lys Tyr Thr Thr Thr Ser Ser Ser Ser Arg Lys
545 550 555 560
Ser Tyr Lys His

<210> 457

<211> 590

<212> PRT

<213> Homo sapiens

<400> 457
Met Ser Arg Gln Ser Ser Val Ser Phe Arg Ser Gly Gly Ser Arg Ser
1 5 10 15
Phe Ser Thr Ala Ser Ala Ile Thr Pro Ser Val Ser Arg Thr Ser Phe
20 25 30
Thr Ser Val Ser Arg Ser Gly Gly Gly Gly Gly Gly Gly Phe Gly Arg
35 40 45
Val Ser Leu Ala Gly Ala Cys Gly Val Gly Gly Tyr Gly Ser Arg Ser
50 55 60
Leu Tyr Asn Leu Gly Gly Ser Lys Arg Ile Ser Ile Ser Thr Arg Gly
65 70 75 80
Gly Ser Phe Arg Asn Arg Phe Gly Ala Gly Ala Gly Gly Gly Tyr Gly
85 90 95
Phe Gly Gly Gly Ala Gly Ser Gly Phe Gly Phe Gly Gly Gly Ala Gly
100 105 110
Gly Gly Phe Gly Leu Gly Gly Gly Ala Gly Phe Gly Gly Gly Phe Gly
115 120 125
Gly Pro Gly Phe Pro Val Cys Pro Pro Gly Gly Ile Gln Glu Val Thr
130 135 140
Val Asn Gln Ser Leu Leu Thr Pro Leu Asn Leu Gln Ile Asp Pro Ser
145 150 155 160
Ile Gln Arg Val Arg Thr Glu Glu Arg Glu Gln Ile Lys Thr Leu Asn
165 170 175
Asn Lys Phe Ala Ser Phe Ile Asp Lys Val Arg Phe Leu Glu Gln Gln
180 185 190
Asn Lys Val Leu Asp Thr Lys Trp Thr Leu Leu Gln Glu Gln Gly Thr
195 200 205
Lys Thr Val Arg Gln Asn Leu Glu Pro Leu Phe Glu Gln Tyr Ile Asn
210 215 220
Asn Leu Arg Arg Gln Leu Asp Ser Ile Val Gly Glu Arg Gly Arg Leu
225 230 235 240
Asp Ser Glu Leu Arg Asn Met Gln Asp Leu Val Glu Asp Phe Lys Asn
245 250 255
Lys Tyr Glu Asp Glu Ile Asn Lys Arg Thr Thr Ala Glu Asn Glu Phe
260 265 270
Val Met Leu Lys Lys Asp Val Asp Ala Ala Tyr Met Asn Lys Val Glu
275 280 285

Leu Glu Ala Lys Val Asp Ala Leu Met Asp Glu Ile Asn Phe Met Lys
290 295 300
Met Phe Phe Asp Ala Glu Leu Ser Gln Met Gln Thr His Val Ser Asp
305 310 315 320
Thr Ser Val Val Leu Ser Met Asp Asn Asn Arg Asn Leu Asp Leu Asp
325 330 335
Ser Ile Ile Ala Glu Val Lys Ala Gln Tyr Glu Glu Ile Ala Asn Arg
340 345 350
Ser Arg Thr Glu Ala Glu Ser Trp Tyr Gln Thr Lys Tyr Glu Glu Leu
355 360 365
Gln Gln Thr Ala Gly Arg His Gly Asp Asp Leu Arg Asn Thr Lys His
370 375 380
Glu Ile Thr Glu Met Asn Arg Met Ile Gln Arg Leu Arg Ala Glu Ile
385 390 395 400
Asp Asn Val Lys Lys Gln Cys Ala Asn Leu Gln Asn Ala Ile Ala Asp
405 410 415
Ala Glu Gln Arg Gly Glu Leu Ala Leu Lys Asp Ala Arg Asn Lys Leu
420 425 430
Ala Glu Leu Glu Glu Ala Leu Gln Lys Ala Lys Gln Asp Met Ala Arg
435 440 445
Leu Leu Arg Glu Tyr Gln Glu Leu Met Asn Thr Lys Leu Ala Leu Asp
450 455 460
Val Glu Ile Ala Thr Tyr Arg Lys Leu Leu Glu Gly Glu Glu Cys Arg
465 470 475 480
Leu Ser Gly Glu Gly Val Gly Pro Val Asn Ile Ser Val Val Thr Ser
485 490 495
Ser Val Ser Ser Gly Tyr Gly Ser Gly Ser Gly Tyr Gly Gly Gly Leu
500 505 510
Gly Gly Gly Leu Gly Gly Gly Leu Gly Gly Gly Leu Ala Gly Gly Ser
515 520 525
Ser Gly Ser Tyr Tyr Ser Ser Ser Ser Gly Gly Val Gly Leu Gly Gly
530 535 540
Gly Leu Ser Val Gly Gly Ser Gly Phe Ser Ala Ser Ser Gly Arg Gly
545 550 555 560
Leu Gly Val Gly Phe Gly Ser Gly Gly Gly Ser Ser Ser Ser Val Lys
565 570 575
Phe Val Ser Thr Thr Ser Ser Ser Arg Lys Ser Phe Lys Ser
580 585 590

<210> 458

<211> 523

<212> PRT

<213> Homo sapiens

<400> 458

Met Ser Arg Gln Phe Thr Cys Lys Ser Gly Ala Ala Ala Lys Gly Gly
1 5 10 15
Phe Ser Gly Cys Ser Ala Val Leu Ser Gly Gly Ser Ser Ser Ser Phe
20 25 30
Arg Ala Gly Ser Lys Gly Leu Ser Gly Gly Phe Gly Ser Arg Ser Leu
35 40 45
Tyr Ser Leu Gly Gly Val Arg Ser Leu Asn Val Ala Ser Gly Ser Gly
50 55 60
Lys Ser Gly Gly Tyr Gly Phe Gly Arg Gly Arg Ala Ser Gly Phe Ala
65 70 75 80
Gly Ser Met Phe Gly Ser Val Ala Leu Gly Pro Val Cys Pro Thr Val
85 90 95
Cys Pro Pro Gly Gly Ile His Gln Val Thr Val Asn Glu Ser Leu Leu
100 105 110
Ala Pro Leu Asn Val Glu Leu Asp Pro Glu Ile Gln Lys Val Arg Ala
115 120 125

Gln Glu Arg Glu Gln Ile Lys Ala Leu Asn Asn Lys Phe Ala Ser Phe
130 135 140
Ile Asp Lys Val Arg Phe Leu Glu Gln Gln Asn Gln Val Leu Glu Thr
145 150 155 160
Lys Trp Glu Leu Leu Gln Gln Leu Asp Leu Asn Asn Cys Lys Asn Asn
165 170 175
Leu Glu Pro Ile Leu Glu Gly Tyr Ile Ser Asn Leu Arg Lys Gln Leu
180 185 190
Glu Thr Leu Ser Gly Asp Arg Val Arg Leu Asp Ser Glu Leu Arg Asn
195 200 205
Val Arg Asp Val Val Glu Asp Tyr Lys Lys Arg Tyr Glu Glu Glu Ile
210 215 220
Asn Lys Arg Thr Ala Ala Glu Asn Glu Phe Val Leu Leu Lys Lys Asp
225 230 235 240
Val Asp Ala Ala Tyr Ala Asn Lys Val Glu Leu Gln Ala Lys Val Glu
245 250 255
Ser Met Asp Gln Glu Ile Lys Phe Phe Arg Cys Leu Phe Glu Ala Glu
260 265 270
Ile Thr Gln Ile Gln Ser His Ile Ser Asp Met Ser Val Ile Leu Ser
275 280 285
Met Asp Asn Asn Arg Asn Leu Asp Leu Asp Ser Ile Ile Asp Glu Val
290 295 300
Arg Thr Gln Tyr Glu Glu Ile Ala Leu Lys Ser Lys Ala Glu Ala Glu
305 310 315 320
Ala Leu Tyr Gln Thr Lys Phe Gln Glu Leu Gln Leu Ala Ala Gly Arg
325 330 335
His Gly Asp Asp Leu Lys Asn Thr Lys Asn Glu Ile Ser Glu Leu Thr
340 345 350
Arg Leu Ile Gln Arg Ile Arg Ser Glu Ile Glu Asn Val Lys Lys Gln
355 360 365
Ala Ser Asn Leu Glu Thr Ala Ile Ala Asp Ala Glu Gln Arg Gly Asp
370 375 380
Asn Ala Leu Lys Asp Ala Arg Ala Lys Leu Asp Glu Leu Glu Gly Ala
385 390 395 400
Leu His Gln Ala Lys Glu Glu Leu Ala Arg Met Leu Arg Glu Tyr Gln
405 410 415
Glu Leu Met Ser Leu Lys Leu Ala Leu Asp Met Glu Ile Ala Thr Tyr
420 425 430
Arg Lys Leu Leu Glu Ser Glu Glu Cys Arg Met Ser Gly Glu Phe Pro
435 440 445
Ser Pro Val Ser Ile Ser Ile Ile Ser Ser Thr Ser Gly Gly Ser Val
450 455 460
Tyr Gly Phe Arg Pro Ser Met Val Ser Gly Gly Tyr Val Ala Asn Ser
465 470 475 480
Ser Asn Cys Ile Ser Gly Val Cys Ser Val Arg Gly Gly Glu Gly Arg
485 490 495
Ser Arg Gly Ser Ala Asn Asp Tyr Lys Asp Thr Leu Gly Lys Gly Ser
500 505 510
Ser Leu Ser Ala Pro Ser Lys Lys Thr Ser Arg
515 520

<210> 459

<211> 529

<212> PRT

<213> Homo sapiens

<400> 459
Met Ser Arg Gln Leu Asn Ile Lys Ser Ser Gly Asp Lys Gly Asn Phe
1 5 10 15
Ser Val His Ser Ala Val Val Pro Arg Lys Ala Val Gly Ser Leu Ala
20 25 30

```
Ser Tyr Cys Ala Ala Gly Arg Gly Ala Gly Ala Gly Phe Gly Ser Arg
        35                  40                  45
Ser Leu Tyr Ser Leu Gly Gly Asn Arg Arg Ile Ser Phe Asn Val Ala
    50                  55                  60
Gly Gly Gly Val Arg Ala Gly Gly Tyr Gly Phe Arg Pro Gly Ser Gly
65                  70                  75                  80
Tyr Gly Gly Gly Arg Ala Ser Gly Phe Ala Gly Ser Met Phe Gly Ser
                85                  90                  95
Val Ala Leu Gly Pro Ala Cys Leu Ser Val Cys Pro Pro Gly Gly Ile
            100                 105                 110
His Gln Val Thr Val Asn Lys Ser Leu Leu Ala Pro Leu Asn Val Glu
        115                 120                 125
Leu Asp Pro Glu Ile Gln Lys Val Arg Ala Gln Glu Arg Glu Gln Ile
    130                 135                 140
Lys Val Leu Asn Asp Lys Phe Ala Ser Phe Ile Asp Lys Val Arg Phe
145                 150                 155                 160
Leu Glu Gln Gln Asn Gln Val Leu Glu Thr Lys Trp Glu Leu Leu Gln
                165                 170                 175
Gln Leu Asp Leu Asn Asn Cys Lys Lys Asn Leu Glu Pro Ile Leu Glu
            180                 185                 190
Gly Tyr Ile Ser Asn Leu Arg Lys Gln Leu Glu Thr Leu Ser Gly Asp
        195                 200                 205
Arg Val Arg Leu Asp Ser Glu Leu Arg Ser Met Arg Asp Leu Val Glu
    210                 215                 220
Asp Tyr Lys Lys Arg Tyr Glu Val Glu Ile Asn Arg Arg Thr Thr Ala
225                 230                 235                 240
Glu Asn Glu Phe Val Val Leu Lys Lys Asp Ala Asp Ala Ala Tyr Ala
                245                 250                 255
Val Lys Val Glu Leu Gln Ala Lys Val Asp Ser Leu Asp Lys Asp Ile
            260                 265                 270
Lys Phe Leu Lys Cys Leu Tyr Asp Ala Glu Ile Ala Gln Ile Gln Thr
        275                 280                 285
His Ala Ser Glu Thr Ser Val Ile Leu Ser Met Asp Asn Asn Arg Asp
    290                 295                 300
Leu Asp Leu Asp Ser Ile Ile Ala Glu Val Arg Met His Tyr Glu Glu
305                 310                 315                 320
Ile Ala Leu Lys Ser Lys Ala Glu Ala Glu Ala Leu Tyr Gln Thr Lys
                325                 330                 335
Ile Gln Glu Leu Gln Leu Ala Ala Ser Arg His Gly Asp Asp Leu Lys
            340                 345                 350
His Thr Arg Ser Glu Met Val Glu Leu Asn Arg Leu Ile Gln Arg Ile
        355                 360                 365
Arg Cys Glu Ile Gly Asn Val Lys Lys Gln Arg Ala Ser Leu Glu Thr
    370                 375                 380
Ala Ile Ala Asp Ala Glu Gln Arg Gly Asp Asn Ala Leu Lys Asp Ala
385                 390                 395                 400
Gln Ala Lys Leu Asp Glu Leu Glu Gly Ala Leu His Gln Ala Lys Glu
                405                 410                 415
Glu Leu Ala Arg Met Leu Arg Glu Tyr Gln Glu Leu Met Ser Leu Lys
            420                 425                 430
Leu Ala Leu Asp Met Glu Ile Ala Thr Tyr Arg Lys Leu Leu Glu Gly
        435                 440                 445
Glu Glu Cys Arg Met Ser Gly Glu Asn Pro Ser Ser Val Ser Ile Ser
    450                 455                 460
Val Ile Ser Ser Ser Ser Tyr Ser Tyr His His Pro Ser Ser Ala Gly
465                 470                 475                 480
Val Asp Leu Gly Ala Ser Ala Val Ala Gly Ser Ser Gly Ser Thr Gln
                485                 490                 495
Ser Gly Gln Thr Lys Thr Thr Glu Ala Arg Gly Gly Asp Leu Lys Asp
            500                 505                 510
Thr Gln Gly Lys Ser Thr Pro Ala Ser Ile Pro Ala Arg Lys Ala Thr
        515                 520                 525
Arg
```

<210> 460

<211> 511

<212> PRT

<213> Homo sapiens

<400> 460

```
Met Ser Arg Gln Leu Thr His Phe Pro Arg Gly Glu Arg Leu Gly Phe
1               5                   10                  15
Ser Gly Cys Ser Ala Val Leu Ser Gly Gly Ile Gly Ser Ser Ser Ala
            20                  25                  30
Ser Phe Arg Ala Arg Val Lys Gly Ser Ala Ser Phe Gly Ser Lys Ser
        35                  40                  45
Leu Ser Cys Leu Gly Gly Ser Arg Ser Leu Ala Leu Ser Ala Ala Ala
    50                  55                  60
Arg Arg Gly Gly Gly Arg Leu Gly Gly Phe Val Gly Thr Ala Phe Gly
65                  70                  75                  80
Ser Ala Gly Leu Gly Pro Lys Cys Pro Ser Val Cys Pro Pro Gly Gly
                85                  90                  95
Ile Pro Gln Val Thr Val Asn Lys Ser Leu Leu Ala Pro Leu Asn Val
            100                 105                 110
Glu Met Asp Pro Glu Ile Gln Arg Val Arg Ala Gln Glu Arg Glu Gln
        115                 120                 125
Ile Lys Ala Leu Asn Asn Lys Phe Ala Ser Phe Ile Asp Lys Val Arg
    130                 135                 140
Phe Leu Glu Gln Gln Asn Gln Val Leu Glu Thr Lys Trp Asn Leu Leu
145                 150                 155                 160
Gln Gln Leu Asp Leu Asn Asn Cys Arg Lys Asn Leu Glu Pro Ile Tyr
                165                 170                 175
Glu Gly Tyr Ile Ser Asn Leu Gln Lys Gln Leu Glu Met Leu Ser Gly
            180                 185                 190
Asp Gly Val Arg Leu Asp Ser Glu Leu Arg Asn Met Gln Asp Leu Val
        195                 200                 205
Glu Asp Tyr Lys Lys Arg Tyr Glu Val Glu Ile Asn Arg Arg Thr Ala
    210                 215                 220
Ala Glu Asn Glu Phe Val Val Leu Lys Lys Asp Val Asp Ala Ala Tyr
225                 230                 235                 240
Met Asn Lys Val Glu Leu Gln Ala Lys Val Asp Ser Leu Thr Asp Glu
                245                 250                 255
Ile Lys Phe Phe Lys Cys Leu Tyr Glu Gly Glu Ile Thr Gln Ile Gln
            260                 265                 270
Ser His Ile Ser Asp Thr Ser Ile Val Leu Ser Met Asp Asn Asn Arg
        275                 280                 285
Asp Leu Asp Leu Asp Ser Ile Ile Ala Glu Val Arg Ala Gln Tyr Glu
    290                 295                 300
Glu Ile Ala Leu Lys Ser Lys Ala Glu Ala Glu Thr Leu Tyr Gln Thr
305                 310                 315                 320
Lys Ile Gln Glu Leu Gln Val Thr Ala Gly Gln His Gly Asp Asp Leu
                325                 330                 335
Lys Leu Thr Lys Ala Glu Ile Ser Glu Leu Asn Arg Leu Ile Gln Arg
            340                 345                 350
Ile Arg Ser Glu Ile Gly Asn Val Lys Lys Gln Cys Ala Asp Leu Glu
        355                 360                 365
Thr Ala Ile Ala Asp Ala Glu Gln Arg Gly Asp Cys Ala Leu Lys Asp
    370                 375                 380
Ala Arg Ala Lys Leu Asp Glu Leu Glu Gly Ala Leu His Gln Ala Lys
385                 390                 395                 400
Glu Glu Leu Ala Arg Met Leu Arg Glu Tyr Gln Glu Leu Val Ser Leu
                405                 410                 415
Lys Leu Ala Leu Asp Met Glu Ile Ala Thr Tyr Arg Lys Leu Leu Glu
            420                 425                 430
Ser Glu Glu Cys Arg Met Ser Gly Glu Tyr Pro Asn Ser Val Ser Ile
        435                 440                 445
```

Ser Val Ile Ser Ser Thr Asn Ala Gly Ala Gly Gly Ala Gly Phe Ser
450 455 460
Met Gly Phe Gly Ala Ser Ser Ser Tyr Ser Tyr Lys Thr Ala Ala Ala
465 470 475 480
Asp Val Lys Thr Lys Gly Ser Cys Gly Ser Glu Leu Lys Asp Pro Leu
485 490 495
Ala Lys Thr Ser Gly Ser Ser Cys Ala Thr Lys Lys Ala Ser Arg
500 505 510

<210> 461

<211> 540

<212> PRT

<213> Homo sapiens

<400> 461
Met Ser Arg Gln Phe Thr Tyr Lys Ser Gly Ala Ala Ala Lys Gly Gly
1 5 10 15
Phe Ser Gly Cys Ser Ala Val Leu Ser Gly Gly Ser Ser Ser Ser Tyr
20 25 30
Arg Ala Gly Gly Lys Gly Leu Ser Gly Gly Phe Ser Ser Arg Ser Leu
35 40 45
Tyr Ser Leu Gly Gly Ala Arg Ser Ile Ser Phe Asn Val Ala Ser Gly
50 55 60
Ser Gly Trp Ala Gly Gly Tyr Gly Phe Gly Arg Gly Arg Ala Ser Gly
65 70 75 80
Phe Ala Gly Ser Met Phe Gly Ser Val Ala Leu Gly Ser Val Cys Pro
85 90 95
Ser Leu Cys Pro Pro Gly Gly Ile His Gln Val Thr Ile Asn Lys Ser
100 105 110
Leu Leu Ala Pro Leu Asn Val Glu Leu Asp Pro Glu Ile Gln Lys Val
115 120 125
Arg Ala Gln Glu Arg Glu Gln Ile Lys Val Leu Asn Asn Lys Phe Ala
130 135 140
Ser Phe Ile Asp Lys Val Arg Phe Leu Glu Gln Gln Asn Gln Val Leu
145 150 155 160
Glu Thr Lys Trp Glu Leu Leu Gln Gln Leu Asp Leu Asn Asn Cys Lys
165 170 175
Asn Asn Leu Glu Pro Ile Leu Glu Gly Tyr Ile Ser Asn Leu Arg Lys
180 185 190
Gln Leu Glu Thr Leu Ser Gly Asp Arg Val Arg Leu Asp Ser Glu Leu
195 200 205
Arg Ser Val Arg Glu Val Val Glu Asp Tyr Lys Lys Arg Tyr Glu Glu
210 215 220
Glu Ile Asn Lys Arg Thr Thr Ala Glu Asn Glu Phe Val Val Leu Lys
225 230 235 240
Lys Asp Val Asp Ala Ala Tyr Thr Ser Lys Val Glu Leu Gln Ala Lys
245 250 255
Val Asp Ala Leu Asp Gly Glu Ile Lys Phe Phe Lys Cys Leu Tyr Glu
260 265 270
Gly Glu Thr Ala Gln Ile Gln Ser His Ile Ser Asp Thr Ser Ile Ile
275 280 285
Leu Ser Met Asp Asn Asn Arg Asn Leu Asp Leu Asp Ser Ile Ile Ala
290 295 300
Glu Val Arg Ala Gln Tyr Glu Glu Ile Ala Arg Lys Ser Lys Ala Glu
305 310 315 320
Ala Glu Ala Leu Tyr Gln Thr Lys Phe Gln Glu Leu Gln Leu Ala Ala
325 330 335
Gly Arg His Gly Asp Asp Leu Lys His Thr Lys Asn Glu Ile Ser Glu
340 345 350
Leu Thr Arg Leu Ile Gln Arg Leu Arg Ser Glu Ile Glu Ser Val Lys
355 360 365

Lys Gln Cys Ala Asn Leu Glu Thr Ala Ile Ala Asp Ala Glu Gln Arg
370 375 380
Gly Asp Cys Ala Leu Lys Asp Ala Arg Ala Lys Leu Asp Glu Leu Glu
385 390 395 400
Gly Ala Leu Gln Gln Ala Lys Glu Glu Leu Ala Arg Met Leu Arg Glu
405 410 415
Tyr Gln Glu Leu Leu Ser Val Lys Leu Ser Leu Asp Ile Glu Ile Ala
420 425 430
Thr Tyr Arg Lys Leu Leu Glu Gly Glu Glu Cys Arg Met Ser Gly Glu
435 440 445
Tyr Thr Asn Ser Val Ser Ile Ser Val Ile Asn Ser Ser Met Ala Gly
450 455 460
Met Ala Gly Thr Gly Ala Gly Phe Gly Phe Ser Asn Ala Gly Thr Tyr
465 470 475 480
Gly Tyr Trp Pro Ser Ser Val Ser Gly Gly Tyr Ser Met Leu Pro Gly
485 490 495
Gly Cys Val Thr Gly Ser Gly Asn Cys Ser Pro Arg Gly Glu Ala Arg
500 505 510
Thr Arg Leu Gly Ser Ala Ser Glu Phe Arg Asp Ser Gln Gly Lys Thr
515 520 525
Leu Ala Leu Ser Ser Pro Thr Lys Lys Thr Met Arg
530 535 540

<210> 462

<211> 645

<212> PRT

<213> Homo sapiens

<400> 462

Met Ser Cys Gln Ile Ser Cys Lys Ser Arg Gly Arg Gly Gly Gly Gly
1 5 10 15
Gly Gly Phe Arg Gly Phe Ser Ser Gly Ser Ala Val Val Ser Gly Gly
20 25 30
Ser Arg Arg Ser Thr Ser Ser Phe Ser Cys Leu Ser Arg His Gly Gly
35 40 45
Gly Gly Gly Gly Phe Gly Gly Gly Gly Phe Gly Ser Arg Ser Leu Val
50 55 60
Gly Leu Gly Gly Thr Lys Ser Ile Ser Ile Ser Val Ala Gly Gly Gly
65 70 75 80
Gly Gly Phe Gly Ala Ala Gly Gly Phe Gly Gly Arg Gly Gly Gly Phe
85 90 95
Gly Gly Gly Ser Gly Phe Gly Gly Gly Ser Gly Phe Gly Gly Gly Ser
100 105 110
Gly Phe Ser Gly Gly Gly Phe Gly Gly Gly Gly Phe Gly Gly Gly Arg
115 120 125
Phe Gly Gly Phe Gly Gly Pro Gly Gly Val Gly Gly Leu Gly Gly Pro
130 135 140
Gly Gly Phe Gly Pro Gly Gly Tyr Pro Gly Gly Ile His Glu Val Ser
145 150 155 160
Val Asn Gln Ser Leu Leu Gln Pro Leu Asn Val Lys Val Asp Pro Glu
165 170 175
Ile Gln Asn Val Lys Ala Gln Glu Arg Glu Gln Ile Lys Thr Leu Asn
180 185 190
Asn Lys Phe Ala Ser Phe Ile Asp Lys Val Arg Phe Leu Glu Gln Gln
195 200 205
Asn Gln Val Leu Gln Thr Lys Trp Glu Leu Leu Gln Gln Met Asn Val
210 215 220
Gly Thr Arg Pro Ile Asn Leu Glu Pro Ile Phe Gln Gly Tyr Ile Asp
225 230 235 240
Ser Leu Lys Arg Tyr Leu Asp Gly Leu Thr Ala Glu Arg Thr Ser Gln
245 250 255

Asn Ser Glu Leu Asn Asn Met Gln Asp Leu Val Glu Asp Tyr Lys Lys
260 265 270
Lys Tyr Glu Asp Glu Ile Asn Lys Arg Thr Ala Ala Glu Asn Asp Phe
275 280 285
Val Thr Leu Lys Lys Asp Val Asp Asn Ala Tyr Met Ile Lys Val Glu
290 295 300
Leu Gln Ser Lys Val Asp Leu Leu Asn Gln Glu Ile Glu Phe Leu Lys
305 310 315 320
Val Leu Tyr Asp Ala Glu Ile Ser Gln Ile His Gln Ser Val Thr Asp
325 330 335
Thr Asn Val Ile Leu Ser Met Asp Asn Ser Arg Asn Leu Asp Leu Asp
340 345 350
Ser Ile Ile Ala Glu Val Lys Ala Gln Tyr Glu Glu Ile Ala Gln Arg
355 360 365
Ser Lys Glu Glu Ala Glu Ala Leu Tyr His Ser Lys Tyr Glu Glu Leu
370 375 380
Gln Val Thr Val Gly Arg His Gly Asp Ser Leu Lys Glu Ile Lys Ile
385 390 395 400
Glu Ile Ser Glu Leu Asn Arg Val Ile Gln Arg Leu Gln Gly Glu Ile
405 410 415
Ala His Val Lys Lys Gln Cys Lys Asn Val Gln Asp Ala Ile Ala Asp
420 425 430
Ala Glu Gln Arg Gly Glu His Ala Leu Lys Asp Ala Arg Asn Lys Leu
435 440 445
Asn Asp Leu Glu Glu Ala Leu Gln Gln Ala Lys Glu Asp Leu Ala Arg
450 455 460
Leu Leu Arg Asp Tyr Gln Glu Leu Met Asn Val Lys Leu Ala Leu Asp
465 470 475 480
Val Glu Ile Ala Thr Tyr Arg Lys Leu Leu Glu Gly Glu Glu Cys Arg
485 490 495
Met Ser Gly Asp Leu Ser Ser Asn Val Thr Val Ser Val Thr Ser Ser
500 505 510
Thr Ile Ser Ser Asn Val Ala Ser Lys Ala Ala Phe Gly Gly Ser Gly
515 520 525
Gly Arg Gly Ser Ser Ser Gly Gly Gly Tyr Ser Ser Gly Ser Ser Ser
530 535 540
Tyr Gly Ser Gly Gly Arg Gln Ser Gly Ser Arg Gly Gly Ser Gly Gly
545 550 555 560
Gly Gly Ser Ile Ser Gly Gly Gly Tyr Gly Ser Gly Gly Gly Ser Gly
565 570 575
Gly Arg Tyr Gly Ser Gly Gly Gly Ser Lys Gly Gly Ser Ile Ser Gly
580 585 590
Gly Gly Tyr Gly Ser Gly Gly Gly Lys His Ser Ser Gly Gly Gly Ser
595 600 605
Arg Gly Gly Ser Ser Ser Gly Gly Gly Tyr Gly Ser Gly Gly Gly Gly
610 615 620
Ser Ser Ser Val Lys Gly Ser Ser Gly Glu Ala Phe Gly Ser Ser Val
625 630 635 640
Thr Phe Ser Phe Arg
645

<210> 463

<211> 644

<212> PRT

<213> Homo sapiens

<400> 463
Met Ser Arg Gln Phe Ser Ser Arg Ser Gly Tyr Arg Ser Gly Gly Gly
1 5 10 15
Phe Ser Ser Gly Ser Ala Gly Ile Ile Asn Tyr Gln Arg Arg Thr Thr
20 25 30

Ser Ser Ser Thr Arg Arg Ser Gly Gly Gly Gly Gly Arg Phe Ser Ser
35 40 45
Cys Gly Gly Gly Gly Gly Ser Phe Gly Ala Gly Gly Gly Phe Gly Ser
50 55 60
Arg Ser Leu Val Asn Leu Gly Gly Ser Lys Ser Ile Ser Ile Ser Val
65 70 75 80
Ala Arg Gly Gly Gly Arg Gly Ser Gly Phe Gly Gly Gly Tyr Gly Gly
85 90 95
Gly Gly Phe Gly Gly Gly Gly Phe Gly Gly Gly Gly Phe Gly Gly Gly
100 105 110
Gly Ile Gly Gly Gly Gly Phe Gly Gly Phe Gly Ser Gly Gly Gly Gly
115 120 125
Phe Gly Gly Gly Gly Phe Gly Gly Gly Gly Tyr Gly Gly Gly Tyr Gly
130 135 140
Pro Val Cys Pro Pro Gly Gly Ile Gln Glu Val Thr Ile Asn Gln Ser
145 150 155 160
Leu Leu Gln Pro Leu Asn Val Glu Ile Asp Pro Glu Ile Gln Lys Val
165 170 175
Lys Ser Arg Glu Arg Glu Gln Ile Lys Ser Leu Asn Asn Gln Phe Ala
180 185 190
Ser Phe Ile Asp Lys Val Arg Phe Leu Glu Gln Gln Asn Gln Val Leu
195 200 205
Gln Thr Lys Trp Glu Leu Leu Gln Gln Val Asp Thr Ser Thr Arg Thr
210 215 220
His Asn Leu Glu Pro Tyr Phe Glu Ser Phe Ile Asn Asn Leu Arg Arg
225 230 235 240
Arg Val Asp Gln Leu Lys Ser Asp Gln Ser Arg Leu Asp Ser Glu Leu
245 250 255
Lys Asn Met Gln Asp Met Val Glu Asp Tyr Arg Asn Lys Tyr Glu Asp
260 265 270
Glu Ile Asn Lys Arg Thr Asn Ala Glu Asn Glu Phe Val Thr Ile Lys
275 280 285
Lys Asp Val Asp Gly Ala Tyr Met Thr Lys Val Asp Leu Gln Ala Lys
290 295 300
Leu Asp Asn Leu Gln Gln Glu Ile Asp Phe Leu Thr Ala Leu Tyr Gln
305 310 315 320
Ala Glu Leu Ser Gln Met Gln Thr Gln Ile Ser Glu Thr Asn Val Ile
325 330 335
Leu Ser Met Asp Asn Asn Arg Ser Leu Asp Leu Asp Ser Ile Ile Ala
340 345 350
Glu Val Lys Ala Gln Tyr Glu Asp Ile Ala Gln Lys Ser Lys Ala Glu
355 360 365
Ala Glu Ser Leu Tyr Gln Ser Lys Tyr Glu Glu Leu Gln Ile Thr Ala
370 375 380
Gly Arg His Gly Asp Ser Val Arg Asn Ser Lys Ile Glu Ile Ser Glu
385 390 395 400
Leu Asn Arg Val Ile Gln Arg Leu Arg Ser Glu Ile Asp Asn Val Lys
405 410 415
Lys Gln Ile Ser Asn Leu Gln Gln Ser Ile Ser Asp Ala Glu Gln Arg
420 425 430
Gly Glu Asn Ala Leu Lys Asp Ala Lys Asn Lys Leu Asn Asp Leu Glu
435 440 445
Asp Ala Leu Gln Gln Ala Lys Glu Asp Leu Ala Arg Leu Leu Arg Asp
450 455 460
Tyr Gln Glu Leu Met Asn Thr Lys Leu Ala Leu Asp Leu Glu Ile Ala
465 470 475 480
Thr Tyr Arg Thr Leu Leu Glu Gly Glu Glu Ser Arg Met Ser Gly Glu
485 490 495
Cys Ala Pro Asn Val Ser Val Ser Val Ser Thr Ser His Thr Thr Ile
500 505 510
Ser Gly Gly Gly Ser Arg Gly Gly Gly Gly Gly Gly Tyr Gly Ser Gly
515 520 525
Gly Ser Ser Tyr Gly Ser Gly Gly Gly Ser Tyr Gly Ser Gly Gly Gly
530 535 540
Gly Gly Gly Gly Arg Gly Ser Tyr Gly Ser Gly Gly Ser Ser Tyr Gly
545 550 555 560

Ser Gly Gly Gly Ser Tyr Gly Ser Gly Gly Gly Gly Gly Gly His Gly
565 570 575
Ser Tyr Gly Ser Gly Ser Ser Ser Gly Gly Tyr Arg Gly Gly Ser Gly
580 585 590
Gly Gly Gly Gly Gly Ser Ser Gly Gly Arg Gly Ser Gly Gly Gly Ser
595 600 605
Ser Gly Gly Ser Ile Gly Gly Arg Gly Ser Ser Ser Gly Gly Val Lys
610 615 620
Ser Ser Gly Gly Ser Ser Ser Val Arg Phe Val Ser Thr Thr Tyr Ser
625 630 635 640
Gly Val Thr Arg

<210> 464

<211> 629

<212> PRT

<213> Homo sapiens

<400> 464
Met Ser Arg Gln Ala Ser Lys Thr Ser Gly Gly Gly Ser Gln Gly Phe
1 5 10 15
Ser Gly Arg Ser Ala Val Val Ser Gly Ser Ser Arg Met Ser Cys Val
20 25 30
Ala His Ser Gly Gly Ala Gly Gly Gly Ala Tyr Gly Phe Arg Ser Gly
35 40 45
Ala Gly Gly Phe Gly Ser Arg Ser Leu Tyr Asn Leu Gly Gly Asn Lys
50 55 60
Ser Ile Ser Ile Ser Val Ala Ala Gly Gly Ser Arg Ala Gly Gly Phe
65 70 75 80
Gly Gly Gly Arg Ser Ser Cys Ala Phe Ala Gly Gly Tyr Gly Gly Gly
85 90 95
Phe Gly Ser Gly Tyr Gly Gly Gly Phe Gly Gly Gly Phe Gly Gly Gly
100 105 110
Arg Gly Met Gly Gly Gly Phe Gly Gly Ala Gly Gly Phe Gly Gly Ala
115 120 125
Gly Gly Phe Gly Gly Ala Gly Gly Phe Gly Gly Pro Gly Gly Phe Gly
130 135 140
Gly Ser Gly Gly Phe Gly Gly Pro Gly Ser Leu Gly Ser Pro Gly Gly
145 150 155 160
Phe Ala Pro Gly Gly Phe Pro Gly Gly Ile Gln Glu Val Thr Thr Asn
165 170 175
Gln Ser Leu Leu Gln Pro Leu Lys Val Glu Thr Asp Pro Gln Ile Gly
180 185 190
Gln Val Lys Ala Gln Glu Arg Glu Gln Ile Lys Thr Leu Asn Asn Lys
195 200 205
Phe Ala Ser Phe Ile Asp Lys Val Arg Phe Leu Glu Gln Gln Asn Lys
210 215 220
Val Leu Glu Thr Lys Trp Asn Leu Leu Gln Gln Gln Gly Thr Ser Ser
225 230 235 240
Ile Ser Gly Thr Asn Asn Leu Glu Pro Leu Phe Glu Asn His Ile Asn
245 250 255
Tyr Leu Arg Ser Tyr Leu Asp Asn Ile Leu Gly Glu Arg Gly Arg Leu
260 265 270
Asp Ser Glu Leu Lys Asn Met Glu Asp Leu Val Glu Asp Phe Lys Lys
275 280 285
Lys Tyr Glu Asp Glu Ile Asn Lys Arg Thr Ala Ala Glu Asn Glu Phe
290 295 300
Val Thr Leu Lys Lys Asp Val Asp Ser Ala Tyr Met Asn Lys Val Glu
305 310 315 320
Leu Gln Ala Lys Val Asp Ala Leu Ile Asp Glu Ile Asp Phe Leu Arg
325 330 335

Thr Leu Tyr Asp Ala Glu Leu Ser Gln Met Gln Ser His Ile Ser Asp
340 345 350
Thr Ser Val Val Leu Ser Met Asp Asn Asn Arg Ser Leu Asp Leu Asp
355 360 365
Ser Ile Ile Ala Glu Val Gly Ala Gln Tyr Glu Asp Ile Ala Gln Arg
370 375 380
Ser Lys Ala Glu Ala Glu Ala Leu Tyr Gln Thr Lys Leu Gly Glu Leu
385 390 395 400
Gln Thr Thr Ala Gly Arg His Gly Asp Asp Leu Arg Asn Thr Lys Ser
405 410 415
Glu Ile Ile Glu Leu Asn Arg Met Ile Gln Arg Leu Arg Ala Glu Ile
420 425 430
Glu Gly Val Lys Lys Gln Asn Ala Asn Leu Gln Thr Ala Ile Ala Gln
435 440 445
Ala Glu Gln His Gly Glu Met Ala Leu Lys Asp Ala Asn Ala Lys Leu
450 455 460
Gln Glu Leu Gln Ala Ala Leu Gln Gln Ala Lys Asp Asp Leu Ala Arg
465 470 475 480
Leu Leu Arg Asp Tyr Gln Glu Leu Met Asn Val Lys Leu Ala Leu Asp
485 490 495
Val Glu Ile Ala Thr Tyr Arg Lys Leu Leu Glu Gly Glu Glu Tyr Ser
500 505 510
Arg Met Ser Gly Glu Cys Pro Ser Ala Val Ser Ile Ser Val Val Ser
515 520 525
Ser Ser Thr Thr Ser Ala Ser Ala Gly Gly Tyr Gly Gly Gly Tyr Gly
530 535 540
Gly Gly Met Gly Gly Gly Leu Gly Gly Gly Phe Ser Ala Gly Gly Gly
545 550 555 560
Ser Gly Ile Gly Phe Gly Arg Gly Gly Gly Gly Gly Ile Gly Gly Gly
565 570 575
Phe Gly Gly Gly Thr Ser Gly Phe Ser Gly Gly Ser Gly Phe Gly Ser
580 585 590
Ile Ser Gly Ala Arg Tyr Gly Val Ser Gly Gly Gly Phe Ser Ser Ala
595 600 605
Ser Asn Arg Gly Gly Ser Ile Lys Phe Ser Gln Ser Ser Gln Ser Ser
610 615 620
Gln Arg Tyr Ser Arg
625

<210> 465

<211> 534

<212> PRT

<213> Homo sapiens

<400> 465

Met Ile Ala Arg Gln Gln Cys Val Arg Gly Gly Pro Arg Gly Phe Ser
1 5 10 15
Cys Gly Ser Ala Ile Val Gly Gly Gly Lys Arg Gly Ala Phe Ser Ser
20 25 30
Val Ser Met Ser Gly Gly Ala Gly Arg Cys Ser Ser Gly Gly Phe Gly
35 40 45
Ser Arg Ser Leu Tyr Asn Leu Arg Gly Asn Lys Ser Ile Ser Met Ser
50 55 60
Val Ala Gly Ser Arg Gln Gly Ala Cys Phe Gly Gly Ala Gly Gly Phe
65 70 75 80
Gly Thr Gly Gly Phe Gly Ala Gly Gly Phe Gly Ala Gly Phe Gly Thr
85 90 95
Gly Gly Phe Gly Gly Gly Phe Gly Gly Ser Phe Ser Gly Lys Gly Gly
100 105 110
Pro Gly Phe Pro Val Cys Pro Ala Gly Gly Ile Gln Glu Val Thr Ile
115 120 125

Asn Gln Ser Leu Leu Thr Pro Leu His Val Glu Ile Asp Pro Glu Ile
130 135 140
Gln Lys Val Arg Thr Glu Glu Arg Glu Gln Ile Lys Leu Leu Asn Asn
145 150 155 160
Lys Phe Ala Ser Phe Ile Asp Lys Val Gln Phe Leu Glu Gln Gln Asn
165 170 175
Lys Val Leu Glu Thr Lys Trp Asn Leu Leu Gln Gln Gln Thr Thr Thr
180 185 190
Thr Ser Ser Lys Asn Leu Glu Pro Leu Phe Glu Thr Tyr Leu Ser Val
195 200 205
Leu Arg Lys Gln Leu Asp Thr Leu Gly Asn Asp Lys Gly Arg Leu Gln
210 215 220
Ser Glu Leu Lys Thr Met Gln Asp Ser Val Glu Asp Phe Lys Thr Lys
225 230 235 240
Tyr Glu Glu Glu Ile Asn Lys Arg Thr Ala Ala Glu Asn Asp Phe Val
245 250 255
Val Leu Lys Lys Asp Val Asp Ala Ala Tyr Leu Asn Lys Val Glu Leu
260 265 270
Glu Ala Lys Val Asp Ser Leu Asn Asp Glu Ile Asn Phe Leu Lys Val
275 280 285
Leu Tyr Asp Ala Glu Leu Ser Gln Met Gln Thr His Val Ser Asp Thr
290 295 300
Ser Val Val Leu Ser Met Asp Asn Asn Arg Asn Leu Asp Leu Asp Ser
305 310 315 320
Ile Ile Ala Glu Val Arg Ala Gln Tyr Glu Glu Ile Ala Gln Arg Ser
325 330 335
Lys Ala Glu Ala Glu Ala Leu Tyr Gln Thr Lys Val Gln Gln Leu Gln
340 345 350
Ile Ser Val Asp Gln His Gly Asp Asn Leu Lys Asn Thr Lys Ser Glu
355 360 365
Ile Ala Glu Leu Asn Arg Met Ile Gln Arg Leu Arg Ala Glu Ile Glu
370 375 380
Asn Ile Lys Lys Gln Cys Gln Thr Leu Gln Val Ser Val Ala Asp Ala
385 390 395 400
Glu Gln Arg Gly Glu Asn Ala Leu Lys Asp Ala His Ser Lys Arg Val
405 410 415
Glu Leu Glu Ala Ala Leu Gln Gln Ala Lys Glu Glu Leu Ala Arg Met
420 425 430
Leu Arg Glu Tyr Gln Glu Leu Met Ser Val Lys Leu Ala Leu Asp Ile
435 440 445
Glu Ile Ala Thr Tyr Arg Lys Leu Leu Glu Gly Glu Glu Tyr Arg Met
450 455 460
Ser Gly Glu Cys Gln Ser Ala Val Ser Ile Ser Val Val Ser Gly Ser
465 470 475 480
Thr Ser Thr Gly Gly Ile Ser Gly Gly Leu Gly Ser Gly Ser Gly Phe
485 490 495
Gly Leu Ser Ser Gly Phe Gly Ser Gly Ser Gly Ser Gly Phe Gly Phe
500 505 510
Gly Gly Ser Val Ser Gly Ser Ser Ser Ser Lys Ile Ile Ser Thr Thr
515 520 525
Thr Leu Asn Lys Arg Arg
530

<210> 466

<211> 483

<212> PRT

<213> Homo sapiens

<400> 466

Met Ser Ile Arg Val Thr Gln Lys Ser Tyr Lys Val Ser Thr Ser Gly
1 5 10 15

Pro Arg Ala Phe Ser Ser Arg Ser Tyr Thr Ser Gly Pro Gly Ser Arg
20 25 30
Ile Ser Ser Ser Ser Phe Ser Arg Val Gly Ser Ser Asn Phe Arg Gly
35 40 45
Gly Leu Gly Gly Gly Tyr Gly Gly Ala Ser Gly Met Gly Gly Ile Thr
50 55 60
Ala Val Thr Val Asn Gln Ser Leu Leu Ser Pro Leu Val Leu Glu Val
65 70 75 80
Asp Pro Asn Ile Gln Ala Val Arg Thr Gln Glu Lys Glu Gln Ile Lys
85 90 95
Thr Leu Asn Asn Lys Phe Ala Ser Phe Ile Asp Lys Val Arg Phe Leu
100 105 110
Glu Gln Gln Asn Lys Met Leu Glu Thr Lys Trp Ser Leu Leu Gln Gln
115 120 125
Gln Lys Thr Ala Arg Ser Asn Met Asp Asn Met Phe Glu Ser Tyr Ile
130 135 140
Asn Asn Leu Arg Arg Gln Leu Glu Thr Leu Gly Gln Glu Lys Leu Lys
145 150 155 160
Leu Glu Ala Glu Leu Gly Asn Met Gln Gly Leu Val Glu Asp Phe Lys
165 170 175
Asn Lys Tyr Glu Asp Glu Ile Asn Lys Arg Thr Glu Met Glu Asn Glu
180 185 190
Phe Val Leu Ile Lys Lys Asp Val Asp Glu Ala Tyr Met Asn Lys Val
195 200 205
Glu Leu Glu Ser Arg Leu Glu Gly Leu Thr Asp Glu Ile Asn Phe Leu
210 215 220
Arg Gln Leu Tyr Glu Glu Ile Arg Glu Leu Gln Ser Gln Ile Ser
225 230 235 240
Asp Thr Ser Val Val Leu Ser Met Asp Asn Ser Arg Ser Leu Asp Met
245 250 255
Asp Ser Ile Ile Ala Glu Val Lys Ala Gln Tyr Glu Asp Ile Ala Asn
260 265 270
Arg Ser Arg Ala Glu Ala Glu Ser Met Tyr Gln Ile Lys Tyr Glu Glu
275 280 285
Leu Gln Ser Leu Ala Gly Lys His Gly Asp Asp Leu Arg Arg Thr Lys
290 295 300
Thr Glu Ile Ser Glu Met Asn Arg Asn Ile Ser Arg Leu Gln Ala Glu
305 310 315 320
Ile Glu Gly Leu Lys Gly Gln Arg Ala Ser Leu Glu Ala Ala Ile Ala
325 330 335
Asp Ala Glu Gln Arg Gly Glu Leu Ala Ile Lys Asp Ala Asn Ala Lys
340 345 350
Leu Ser Glu Leu Glu Ala Ala Leu Gln Arg Ala Lys Gln Asp Met Ala
355 360 365
Arg Gln Leu Arg Glu Tyr Gln Glu Leu Met Asn Val Lys Leu Ala Leu
370 375 380
Asp Ile Glu Ile Ala Thr Tyr Arg Lys Leu Leu Glu Gly Glu Glu Ser
385 390 395 400
Arg Leu Glu Ser Gly Met Gln Asn Met Ser Ile His Thr Lys Thr Thr
405 410 415
Ser Gly Tyr Ala Gly Gly Leu Ser Ser Ala Tyr Gly Gly Leu Thr Ser
420 425 430
Pro Gly Leu Ser Tyr Ser Leu Gly Ser Ser Phe Gly Ser Gly Ala Gly
435 440 445
Ser Ser Ser Phe Ser Arg Thr Ser Ser Ser Arg Ala Val Val Val Lys
450 455 460
Lys Ile Glu Thr Arg Asp Gly Lys Leu Val Ser Glu Ser Ser Asp Val
465 470 475 480
Leu Pro Lys

<210> 467

<211> 430

<212> PRT

<213> Homo sapiens

<400> 467
Met Ser Phe Thr Thr Arg Ser Thr Phe Ser Thr Asn Tyr Arg Ser Leu
1 5 10 15
Gly Ser Val Gln Ala Pro Ser Tyr Gly Ala Arg Pro Val Ser Ser Ala
20 25 30
Ala Ser Val Tyr Ala Gly Ala Gly Gly Ser Gly Ser Arg Ile Ser Val
35 40 45
Ser Arg Ser Thr Ser Phe Arg Gly Gly Met Gly Ser Gly Gly Leu Ala
50 55 60
Thr Gly Ile Ala Gly Gly Leu Ala Gly Met Gly Gly Ile Gln Asn Glu
65 70 75 80
Lys Glu Thr Met Gln Ser Leu Asn Asp Arg Leu Ala Ser Tyr Leu Asp
85 90 95
Arg Val Arg Ser Leu Glu Thr Glu Asn Arg Arg Leu Glu Ser Lys Ile
100 105 110
Arg Glu His Leu Glu Lys Lys Gly Pro Gln Val Arg Asp Trp Ser His
115 120 125
Tyr Phe Lys Ile Ile Glu Asp Leu Arg Ala Gln Ile Phe Ala Asn Thr
130 135 140
Val Asp Asn Ala Arg Ile Val Leu Gln Ile Asp Asn Ala Arg Leu Ala
145 150 155 160
Ala Asp Asp Phe Arg Val Lys Tyr Glu Thr Glu Leu Ala Met Arg Gln
165 170 175
Ser Val Glu Asn Asp Ile His Gly Leu Arg Lys Val Ile Asp Asp Thr
180 185 190
Asn Ile Thr Arg Leu Gln Leu Glu Thr Glu Ile Glu Ala Leu Lys Glu
195 200 205
Glu Leu Leu Phe Met Lys Lys Asn His Glu Glu Glu Val Lys Gly Leu
210 215 220
Gln Ala Gln Ile Ala Ser Ser Gly Leu Thr Val Glu Val Asp Ala Pro
225 230 235 240
Lys Ser Gln Asp Leu Ala Lys Ile Met Ala Asp Ile Arg Ala Gln Tyr
245 250 255
Asp Glu Leu Ala Arg Lys Asn Arg Glu Glu Leu Asp Lys Tyr Trp Ser
260 265 270
Gln Gln Ile Glu Glu Ser Thr Thr Val Val Thr Thr Gln Ser Ala Glu
275 280 285
Val Gly Ala Ala Glu Thr Thr Leu Thr Glu Leu Arg Arg Thr Val Gln
290 295 300
Ser Leu Glu Ile Asp Leu Asp Ser Met Arg Asn Leu Lys Ala Ser Leu
305 310 315 320
Glu Asn Ser Leu Arg Glu Val Glu Ala Arg Tyr Ala Leu Gln Met Glu
325 330 335
Gln Leu Asn Gly Ile Leu Leu His Leu Glu Ser Glu Leu Ala Gln Thr
340 345 350
Arg Ala Glu Gly Gln Arg Gln Ala Gln Glu Tyr Glu Ala Leu Leu Asn
355 360 365
Ile Lys Val Lys Leu Glu Ala Glu Ile Ala Thr Tyr Arg Arg Leu Leu
370 375 380
Glu Asp Gly Glu Asp Phe Asn Leu Gly Asp Ala Leu Asp Ser Ser Asn
385 390 395 400
Ser Met Gln Thr Ile Gln Lys Thr Thr Thr Arg Arg Ile Val Asp Gly
405 410 415
Lys Val Val Ser Glu Thr Asn Asp Thr Lys Val Leu Arg His
420 425 430

<210> 468

<211> 392

<212> PRT

<213> Homo sapiens

<400> 468

Met Val Ala Arg Val Gly Leu Leu Leu Arg Ala Leu Gln Leu Leu Leu
1 5 10 15
Trp Gly His Leu Asp Ala Gln Pro Ala Glu Arg Gly Gly Gln Glu Leu
20 25 30
Arg Lys Glu Ala Glu Ala Phe Leu Glu Lys Tyr Gly Tyr Leu Asn Glu
35 40 45
Gln Val Pro Lys Ala Pro Thr Ser Thr Arg Phe Ser Asp Ala Ile Arg
50 55 60
Ala Phe Gln Trp Val Ser Gln Leu Pro Val Ser Gly Val Leu Asp Arg
65 70 75 80
Ala Thr Leu Arg Gln Met Thr Arg Pro Arg Cys Gly Val Thr Asp Thr
85 90 95
Asn Ser Tyr Ala Ala Trp Ala Glu Arg Ile Ser Asp Leu Phe Ala Arg
100 105 110
His Arg Thr Lys Met Arg Arg Lys Lys Arg Phe Ala Lys Gln Gly Asn
115 120 125
Lys Trp Tyr Lys Gln His Leu Ser Tyr Arg Leu Val Asn Trp Pro Glu
130 135 140
His Leu Pro Glu Pro Ala Val Arg Gly Ala Val Arg Ala Ala Phe Gln
145 150 155 160
Leu Trp Ser Asn Val Ser Ala Leu Glu Phe Trp Glu Ala Pro Ala Thr
165 170 175
Gly Pro Ala Asp Ile Arg Leu Thr Phe Phe Gln Gly Asp His Asn Asp
180 185 190
Gly Leu Gly Asn Ala Phe Asp Gly Pro Gly Gly Ala Leu Ala His Ala
195 200 205
Phe Leu Pro Arg Arg Gly Glu Ala His Phe Asp Gln Asp Glu Arg Trp
210 215 220
Ser Leu Ser Arg Arg Arg Gly Arg Asn Leu Phe Val Val Leu Ala His
225 230 235 240
Glu Ile Gly His Thr Leu Gly Leu Thr His Ser Pro Ala Pro Arg Ala
245 250 255
Leu Met Ala Pro Tyr Tyr Lys Arg Leu Gly Arg Asp Ala Leu Leu Ser
260 265 270
Trp Asp Asp Val Leu Ala Val Gln Ser Leu Tyr Gly Lys Pro Leu Gly
275 280 285
Gly Ser Val Ala Val Gln Leu Pro Gly Lys Leu Phe Thr Asp Phe Glu
290 295 300
Thr Trp Asp Ser Tyr Ser Pro Gln Gly Arg Arg Pro Glu Thr Gln Gly
305 310 315 320
Pro Lys Tyr Cys His Ser Ser Phe Asp Ala Ile Thr Val Asp Arg Gln
325 330 335
Gln Gln Leu Tyr Ile Phe Lys Gly Ser His Phe Trp Glu Val Ala Ala
340 345 350
Asp Gly Asn Val Ser Glu Pro Arg Pro Leu Gln Glu Arg Trp Val Gly
355 360 365
Leu Pro Pro Asn Ile Glu Ala Ala Ala Val Ser Leu Asn Asp Gly Asp
370 375 380
Phe Tyr Phe Phe Lys Val Gln Ser
385 390

<210> 469

<211> 851

<212> PRT

<213> Homo sapiens

<400> 469
Met Ala Gln Trp Glu Met Leu Gln Asn Leu Asp Ser Pro Phe Gln Asp
1 5 10 15
Gln Leu His Gln Leu Tyr Ser His Ser Leu Leu Pro Val Asp Ile Arg
20 25 30
Gln Tyr Leu Ala Val Trp Ile Glu Asp Gln Asn Trp Gln Glu Ala Ala
35 40 45
Leu Gly Ser Asp Asp Ser Lys Ala Thr Met Leu Phe Phe His Phe Leu
50 55 60
Asp Gln Leu Asn Tyr Glu Cys Gly Arg Cys Ser Gln Asp Pro Glu Ser
65 70 75 80
Leu Leu Leu Gln His Asn Leu Arg Lys Phe Cys Arg Asp Ile Gln Pro
85 90 95
Phe Ser Gln Asp Pro Thr Gln Leu Ala Glu Met Ile Phe Asn Leu Leu
100 105 110
Leu Glu Glu Lys Arg Ile Leu Ile Gln Ala Gln Arg Ala Gln Leu Glu
115 120 125
Gln Gly Glu Pro Val Leu Glu Thr Pro Val Glu Ser Gln Gln His Glu
130 135 140
Ile Glu Ser Arg Ile Leu Asp Leu Arg Ala Met Met Glu Lys Leu Val
145 150 155 160
Lys Ser Ile Ser Gln Leu Lys Asp Gln Gln Asp Val Phe Cys Phe Arg
165 170 175
Tyr Lys Ile Gln Ala Lys Gly Lys Thr Pro Ser Leu Asp Pro His Gln
180 185 190
Thr Lys Glu Gln Lys Ile Leu Gln Glu Thr Leu Asn Glu Leu Asp Lys
195 200 205
Arg Arg Lys Glu Val Leu Asp Ala Ser Lys Ala Leu Leu Gly Arg Leu
210 215 220
Thr Thr Leu Ile Glu Leu Leu Leu Pro Lys Leu Glu Glu Trp Lys Ala
225 230 235 240
Gln Gln Gln Lys Ala Cys Ile Arg Ala Pro Ile Asp His Gly Leu Glu
245 250 255
Gln Leu Glu Thr Trp Phe Thr Ala Gly Ala Lys Leu Leu Phe His Leu
260 265 270
Arg Gln Leu Leu Lys Glu Leu Lys Gly Leu Ser Cys Leu Val Ser Tyr
275 280 285
Gln Asp Asp Pro Leu Thr Lys Gly Val Asp Leu Arg Asn Ala Gln Val
290 295 300
Thr Glu Leu Leu Gln Arg Leu Leu His Arg Ala Phe Val Val Glu Thr
305 310 315 320
Gln Pro Cys Met Pro Gln Thr Pro His Arg Pro Leu Ile Leu Lys Thr
325 330 335
Gly Ser Lys Phe Thr Val Arg Thr Arg Leu Leu Val Arg Leu Gln Glu
340 345 350
Gly Asn Glu Ser Leu Thr Val Glu Val Ser Ile Asp Arg Asn Pro Pro
355 360 365
Gln Leu Gln Gly Phe Arg Lys Phe Asn Ile Leu Thr Ser Asn Gln Lys
370 375 380
Thr Leu Thr Pro Glu Lys Gly Gln Ser Gln Gly Leu Ile Trp Asp Phe
385 390 395 400
Gly Tyr Leu Thr Leu Val Glu Gln Arg Ser Gly Gly Ser Gly Lys Gly
405 410 415
Ser Asn Lys Gly Pro Leu Gly Val Thr Glu Glu Leu His Ile Ile Ser
420 425 430
Phe Thr Val Lys Tyr Thr Tyr Gln Gly Leu Lys Gln Glu Leu Lys Thr
435 440 445
Asp Thr Leu Pro Val Val Ile Ile Ser Asn Met Asn Gln Leu Ser Ile
450 455 460
Ala Trp Ala Ser Val Leu Trp Phe Asn Leu Leu Ser Pro Asn Leu Gln
465 470 475 480
Asn Gln Gln Phe Phe Ser Asn Pro Pro Lys Ala Pro Trp Ser Leu Leu
485 490 495
Gly Pro Ala Leu Ser Trp Gln Phe Ser Ser Tyr Val Gly Arg Gly Leu
500 505 510

Asn Ser Asp Gln Leu Ser Met Leu Arg Asn Lys Leu Phe Gly Gln Asn
515 520 525
Cys Arg Thr Glu Asp Pro Leu Leu Ser Trp Ala Asp Phe Thr Lys Arg
530 535 540
Glu Ser Pro Pro Gly Lys Leu Pro Phe Trp Thr Trp Leu Asp Lys Ile
545 550 555 560
Leu Glu Leu Val His Asp His Leu Lys Asp Leu Trp Asn Asp Gly Arg
565 570 575
Ile Met Gly Phe Val Ser Arg Ser Gln Glu Arg Arg Leu Leu Lys Lys
580 585 590
Thr Met Ser Gly Thr Phe Leu Leu Arg Phe Ser Glu Ser Ser Glu Gly
595 600 605
Gly Ile Thr Cys Ser Trp Val Glu His Gln Asp Asp Asp Lys Val Leu
610 615 620
Ile Tyr Ser Val Gln Pro Tyr Thr Lys Glu Val Leu Gln Ser Leu Pro
625 630 635 640
Leu Thr Glu Ile Ile Arg His Tyr Gln Leu Leu Thr Glu Glu Asn Ile
645 650 655
Pro Glu Asn Pro Leu Arg Phe Leu Tyr Pro Arg Ile Pro Arg Asp Glu
660 665 670
Ala Phe Gly Cys Tyr Tyr Gln Glu Lys Val Asn Leu Gln Glu Arg Arg
675 680 685
Lys Tyr Leu Lys His Arg Leu Ile Val Val Ser Asn Arg Gln Val Asp
690 695 700
Glu Leu Gln Gln Pro Leu Glu Leu Lys Pro Glu Pro Glu Leu Glu Ser
705 710 715 720
Leu Glu Leu Glu Leu Gly Leu Val Pro Glu Pro Glu Leu Ser Leu Asp
725 730 735
Leu Glu Pro Leu Leu Lys Ala Gly Leu Asp Leu Gly Pro Glu Leu Glu
740 745 750
Ser Val Leu Glu Ser Thr Leu Glu Pro Val Ile Glu Pro Thr Leu Cys
755 760 765
Met Val Ser Gln Thr Val Pro Glu Pro Asp Gln Gly Pro Val Ser Gln
770 775 780
Pro Val Pro Glu Pro Asp Leu Pro Cys Asp Leu Arg His Leu Asn Thr
785 790 795 800
Glu Pro Met Glu Ile Phe Arg Asn Cys Val Lys Ile Glu Glu Ile Met
805 810 815
Pro Asn Gly Asp Pro Leu Leu Ala Gly Gln Asn Thr Val Asp Glu Val
820 825 830
Tyr Val Ser Arg Pro Ser His Phe Tyr Thr Asp Gly Pro Leu Met Pro
835 840 845
Ser Asp Phe
850

<210> 470

<211> 335

<212> PRT

<213> Homo sapiens

<400> 470
Met Gly Lys Val Lys Val Gly Val Asn Gly Phe Gly Arg Ile Gly Arg
1 5 10 15
Leu Val Thr Arg Ala Ala Phe Asn Ser Gly Lys Val Asp Ile Val Ala
20 25 30
Ile Asn Asp Pro Phe Ile Asp Leu Asn Tyr Met Val Tyr Met Phe Gln
35 40 45
Tyr Asp Ser Thr His Gly Lys Phe His Gly Thr Val Lys Ala Glu Asn
50 55 60
Gly Lys Leu Val Ile Asn Gly Asn Pro Ile Thr Ile Phe Gln Glu Arg
65 70 75 80

Asp Pro Ser Lys Ile Lys Trp Gly Asp Ala Gly Ala Glu Tyr Val Val
85 90 95
Glu Ser Thr Gly Val Phe Thr Thr Met Glu Lys Ala Gly Ala His Leu
100 105 110
Gln Gly Gly Ala Lys Arg Val Ile Ile Ser Ala Pro Ser Ala Asp Ala
115 120 125
Pro Met Phe Val Met Gly Val Asn His Glu Lys Tyr Asp Asn Ser Leu
130 135 140
Lys Ile Ile Ser Asn Ala Ser Cys Thr Thr Asn Cys Leu Ala Pro Leu
145 150 155 160
Ala Lys Val Ile His Asp Asn Phe Gly Ile Val Glu Gly Leu Met Thr
165 170 175
Thr Val His Ala Ile Thr Ala Thr Gln Lys Thr Val Asp Gly Pro Ser
180 185 190
Gly Lys Leu Trp Arg Asp Gly Arg Gly Ala Leu Gln Asn Ile Ile Pro
195 200 205
Ala Ser Thr Gly Ala Ala Lys Ala Val Gly Lys Val Ile Pro Glu Leu
210 215 220
Asn Gly Lys Leu Thr Gly Met Ala Phe Arg Val Pro Thr Ala Asn Val
225 230 235 240
Ser Val Val Asp Leu Thr Cys Arg Leu Glu Lys Pro Ala Lys Tyr Asp
245 250 255
Asp Ile Lys Lys Val Val Lys Gln Ala Ser Glu Gly Pro Leu Lys Gly
260 265 270
Ile Leu Gly Tyr Thr Glu His Gln Val Val Ser Ser Asp Phe Asn Ser
275 280 285
Asp Thr His Ser Ser Thr Phe Asp Ala Gly Ala Gly Ile Ala Leu Asn
290 295 300
Asp His Phe Val Lys Leu Ile Ser Trp Tyr Asp Asn Glu Phe Gly Tyr
305 310 315 320
Ser Asn Arg Val Val Asp Leu Met Ala His Met Ala Ser Lys Glu
325 330 335

<210> 471

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> NAP4 Probe
<400> 471
tccgcctcag tcgcctcttt cg 22
<210> 472

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> NAP4 FOR PRIMER
<400> 472
tcggaagggc tccttcaaa 19
<210> 473

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> NAP4 REV PRIMER
<400> 473
caccgttgca gctcttggt 19
<210> 474

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> MRLP45 Probe
<400> 474
ctcccattcc cctcatgcta taaaaagaac tacc 34
<210> 475

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> MRLP45 FOR PRIMER
<400> 475
ggctgctgga agctttgaag 20
<210> 476

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> MRLP45 REV PRIMER
<400> 476
tgagcaggat gggagagaac a 21
<210> 477

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> TCF2 Probe
<400> 477
caaaagctgg ccatggacgc ct 22
<210> 478

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> TCF2 FOR PRIMER
<400> 478
gcaggaagga ggaggcattc 20
<210> 479

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> TCF2 REV PRIMER
<400> 479
caggctgtga gtctggttgg a 21
<210> 480

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> ROK1 Probe
<400> 480
cagctggctt ccattttcct ggcct 25
<210> 481

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> ROK1 FOR PRIMER
<400> 481
tggcaaaact gggttcagag a 21
<210> 482

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> ROK1 REV PRIMER
<400> 482
tcggaccttg tgggatgtg 19
<210> 483

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT1 Probe
<400> 483
ccgccgccta atatgcaaca ttaggg 26
<210> 484

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT1 FOR PRIMER
<400> 484
cgagtattcc aaagctggta tcg 23
<210> 485

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT1 REV PRIMER
<400> 485
atcacagaga gatggccctt atct 24
<210> 486

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT5 Probe
<400> 486
ccgccgccta atatgcaaca ttaggg 26
<210> 487

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT5 FOR PRIMER
<400> 487
cgagtattcc aaagctggta tcg 23
<210> 488

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT5 REV PRIMER
<400> 488
atcacagaga gatggccctt atct 24
<210> 489

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT8 Probe
<400> 489
ccgccgccta atatgcaaca ttaggg 26
<210> 490

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT8 FOR PRIMER

<400> 490
cgagtattcc aaagctggta tcg 23
<210> 491

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT8 REV PRIMER
<400> 491
atcacagaga gatggccctt atct 24
<210> 492

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT9 Probe
<400> 492
ccgccgccta atatgcaaca ttaggg 26
<210> 493

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT9 FOR PRIMER
<400> 493
cgagtattcc aaagctggta tcg 23
<210> 494

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT9 REV PRIMER
<400> 494
atcacagaga gatggccctt atct 24
<210> 495

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT10-2 Probe
<400> 495
ccgccgccta atatgcaaca ttaggg 26
<210> 496

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT10-2 FOR PRIMER
<400> 496
cgagtattcc aaagctggta tcg 23
<210> 497

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT10-2 REV PRIMER
<400> 497
atcacagaga gatggccctt atct 24
<210> 498

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT14 Probe
<400> 498
ccgccgccta atatgcaaca ttaggg 26
<210> 499

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT14 FOR PRIMER
<400> 499
cgagtattcc aaagctggta tcg 23
<210> 500

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT14 REV PRIMER
<400> 500
atcacagaga gatggccctt atct 24
<210> 501

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT18 Probe
<400> 501
ccgccgccta atatgcaaca ttaggg 26
<210> 502

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT18 FOR PRIMER
<400> 502
cgagtattcc aaagctggta tcg 23
<210> 503

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT18 REV PRIMER
<400> 503
atcacagaga gatggccctt atct 24
<210> 504

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT19 Probe
<400> 504
ccgccgccta atatgcaaca ttaggg 26
<210> 505

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT19 FOR PRIMER
<400> 505
cgagtattcc aaagctggta tcg 23
<210> 506

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT19 REV PRIMER
<400> 506
atcacagaga gatggccctt atct 24
<210> 507

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT6a/b Probe
<400> 507
ccgccgccta atatgcaaca ttaggg 26
<210> 508

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT6a/b FOR PRIMER
<400> 508
cgagtattcc aaagctggta tcg 23
<210> 509

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT6a/b REV PRIMER
<400> 509
atcacagaga gatggccctt atct 24
<210> 510

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT20 Probe
<400> 510
tggcgggaat cctatttatc agactctgta attga 35
<210> 511

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> KRT20 FOR PRIMER
<400> 511
gcaagaaatc agccataaga aagc 24
<210> 512

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 512
ttgcagctcc tctgagtaaa acat 24